

**Perzeptorische Untersuchungen  
zur Intonation der Frage im Russischen**

Schriftliche Hausarbeit zur Erlangung des Grades  
eines Magister Artium (M.A.)  
der Philosophischen Fakultät  
der Christian-Albrechts-Universität  
zu Kiel

vorgelegt von  
Tamara Khromovskikh

Kiel, 2003

Referent: Prof. Dr. K. J. Kohler

Korreferent: Prof. J. Harrington

Tag der mündlichen Prüfung:

Zur Vervielfältigung genehmigt: Kiel, den

Dekan: Prof. Dr. A. Meier

## **1. Einleitung**

Intonation, oder Sprechmelodie, ist ein wichtiges Kommunikationsmittel, das in den Sprachen der Welt zum Ausdruck verschiedener grammatischer Erscheinungen, semantisch-pragmatischer Konzepte und expressiver Bedeutungen eingesetzt wird. Diese wichtige Eigenschaft der menschlichen Sprache steht im Mittelpunkt der vorliegenden Arbeit. Im Kapitel 2. gilt es zunächst, einen Überblick über das komplexe Problemfeld des intonatorischen Phänomenbereichs im Licht seiner kommunikativen Bedeutung zu geben. Dort werden auch einige aktuelle Ansätze und Methoden der Intonationsforschung erläutert und somit die Grundlagen für das Verständnis dieser Arbeit geschaffen. Im Unterkapitel 2.2. soll gezeigt werden, daß der intonatorischen Gestaltung von Äußerungen eine hohe Relevanz hinsichtlich der Markierung des Satzmodus zukommt. Das Wesen der Frage (im Gegensatz zur Aussage) und deren formaler Ausdruck in verschiedenen Sprachen der Welt und speziell im Russischen werden ebenfalls diskutiert, wobei insbesondere die Rolle der Intonation berücksichtigt wird.

Die empirische Hauptuntersuchung der vorliegenden Arbeit ist ein Beitrag zur Erforschung der Wahrnehmung von Äußerungsintonation im Russischen, wobei sich das spezielle Augenmerk auf die Satzmodusdifferenzierung zwischen Frage und Aussage richtet. Auf das Problem, welche Formmittel sich auf dem Gebiet der Intonation zur Kodierung der Aussage- vs. Fragemodus im modernen Russisch etabliert haben, soll das theoretische Unterkapitel 3.1. eingehen. Dort wird ein Überblick über die Kategorisierungsansätze gegeben, wie sie aufgrund von Produktionsdaten in der einschlägigen Literatur zum russischen Intonationssystem vorgenommen wurden. Die Darstellung der wichtigsten Ansätze soll die bisher noch lückenhafte Beschreibung der akustischen Unterschiede im Grundfrequenzverlauf von Frage- vs. Aussageäußerungen aufzeigen. Im Unterkapitel 3.2. wird ein Versuch unternommen, die Konturbeschreibung durch fehlende akustische Informationen zu vervollständigen. Das vierte Theoriekapitel dieser Arbeit beschäftigt sich mit Perzeptionsuntersuchungen zum Russischen (4.1.) sowie zu anderen Sprachen (4.2.), in denen ebenfalls ein System intonatorischer Kontraste für die Satzmodusunterscheidung verwendet wird. Auf Grundlage der Ergebnisse aus der aktuellen Produktions- und Perzeptionsforschung und der unter 3.2. dargestellten empirischen Voruntersuchung werden Hypothesen für die durchzuführende Perzeptionsuntersuchung abgeleitet, in der es um die Frage geht, welche akustischen Schlüsselreize (*acoustic cues*) für russische Hörer zum Dekodieren der Sprecherintention relevant sind und wie diese *cues* gewichtet werden. Die

Perzeptionsuntersuchung wird im fünften Kapitel dargestellt. Im sechsten Kapitel werden die empirisch gewonnenen Ergebnisse zur Perzeption von intonatorisch markierten Fragen im Gegensatz zu Aussagen im Russischen diskutiert (Abschnitt 6.1.). Einen besonderen Diskussionspunkt bildet dabei die Universalienforschung (6.2.). Des weiteren wird in der Diskussion sowohl auf die methodischen Probleme des Experiments (6.3.), als auch auf die Möglichkeiten weiterführender Forschung (6.4.) eingegangen. Schließlich gibt das Kapitel 7 eine kurze Zusammenfassung der Ergebnisse der durchgeführten Perzeptionsuntersuchung und einen Ausblick.

Zu den Formalia der vorliegenden Arbeit ist zu sagen, daß alle russischen Äußerungen im Folgenden in lateinischer Schrift angegeben werden. Die hier verwendeten Transliterationsregeln sind dem Anhang A zu entnehmen. Die den Satzakzent tragenden Silben werden stets mit Fettdruck versehen. Sämtliche für die vorliegende Arbeit verwendeten Sprachsignale (produzierte Originale sowie manipulierte Stimuli) befinden sich auf der beiliegenden CD-ROM (im WAV-Format). Unter dem Verzeichnis *ProduktionsDaten* sind dabei die Signale gespeichert, die für die Voruntersuchung (Unterkapitel 3.2.) aufgenommen und ausgewertet wurden. Unter dem Verzeichnis *PerzeptionsSignale* finden sich die für die durchgeführten Perzeptionsexperimente (Kapitel 5.) relevanten Signaldateien. Die vorgenommenen Wort- und Segment-Etikettierungen der Signale in lateinischer Schrift (s. oben) können den begleitenden Dateien im MIX-Format entnommen werden.

## **2. Theoretische Grundlagen**

Im folgenden Kapitel geht es darum, einige theoretische Grundlagen aus den wissenschaftlichen Teildisziplinen der Phonetik und Linguistik zu erläutern, in denen die vorliegende Arbeit anzusiedeln ist. Hierbei wird in erster Linie eine Begriffs- und Ansatzklärung vorgenommen.

### **2.1. Problemfeld der Intonation**

#### **2.1.1. Intonation als kommunikatives Mittel**

Die lautsprachliche Kommunikation stellt in der Regel eine Interaktion zwischen zwei oder mehreren Gesprächspartnern dar, bei der ein Austausch von verschiedenartigen Informationen

stattfindet. Ein kommunikatives Ereignis kann als Zusammenwirken von vier Manifestationsebenen aufgefaßt werden (s. Abb. 2./1.).

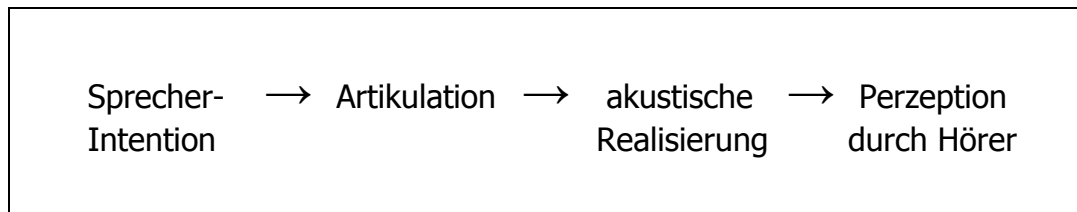


Abb. 2./1.: Vier Stufen der lautsprachlichen Kommunikation (nach Werner und Keller, 1994: 23).

Damit eine korrekte Informationsübermittlung gemäß der Sprecherintention stattfindet und die beabsichtigte Wirkung auf den Perzipienten eintritt, müssen die Informationen entsprechend den für das jeweilige Sprachsystem gültigen formalen Gesetzmäßigkeiten kodiert und realisiert werden.

Zu den kommunikativ bedeutsamen Kodierungssystemen der Sprache gehört neben der segmentellen Ebene die *Intonation*. Die intonatorischen Merkmale der gesprochenen Sprache tragen in erheblichem Maße dazu bei, daß die kommunikativen Sprecherintentionen wie Behauptungen, Stellungnahmen im Gegensatz zu Fragen oder Aufforderungen für den Hörer deutlich werden. Auch emotionale und kognitive Einstellungen des Sprechers dem Hörer und dem Kommunikationsgegenstand gegenüber können durch intonatorische Mittel signalisiert werden. Darüber hinaus dient die Intonation der Verständlichkeit gesprochener Sprache, indem sie zur lexikalischen und syntaktischen Desambiguierung, der Herstellung textueller Kohärenz, aber auch der Strukturierung des Sprechflusses in einzelne kommunikative Einheiten eingesetzt wird. Bedeutsame Phrasenkonstituenten können durch intonatorische Mittel hervorgehoben werden und somit zur Gliederung der Äußerungen in Fokus- vs. Hintergrund-Informationen beitragen.

Die Intonation wird oft als „*an ensemble of pitch variations in the course of utterance*“ definiert (’t Hart, Collier und Cohen, 1990: 10). Da die Intonation nicht zu den Eigenschaften eines einzelnen Segments gehört, sondern auf Äußerungsebene fungiert, zählt sie zu der sogenannten suprasegmentellen (oder prosodischen) Sprachdomäne. Diese wird durch eine Reihe von kontrastierenden Merkmalen über größeren zeitlichen Bereichen konstituiert und umfaßt neben Intonation auch Sprechtempo, Lautheit, Rhythmus, Stimmqualität (vgl. Möbius, 1993).

### **2.1.2. Manifestationsebenen der Intonation**

Das sprachliche Ausdrucksmittel Intonation kann auf der Ebene der Produktion durch einen Sprecher, auf der Ebene der Übertragung durch ein Medium sowie auf der Ebene der Perzeption durch einen Hörer analysiert werden. Auf der **artikulatorischen Manifestationsebene** hängt die Intonation primär mit der Schwingungsrate der sich öffnenden und schließenden Stimmlippen im Larynx zusammen. Diesem physiologischen Prozeß der Stimmerzeugung liegen myoelastische Eigenschaften der laryngalen Muskulatur und aerodynamische Kräfte der oberen Atemwege zugrunde (van den Berg, 1968). Die sprecherspezifischen Abweichungen in der Vibrationsgeschwindigkeit werden durch die individuelle Länge und Masse der Stimmlippen determiniert: Je massiver (länger und dicker) Stimmlippen beschaffen sind, desto langsamer erfolgt eine komplette Schwingungsphase des Öffnens und Schließens der Glottis. Die physiologischen Unterschiede in der Stimmlippenbeschaffenheit unterliegen der Geschlechts- und Altersvariation (vgl. Helfrich, 1979). Die funktional relevanten Änderungen des Schwingungsverhaltens der Stimmlippen, die weitestgehend der Sprecherkontrolle unterliegen und in verschiedenen Melodieverläufen resultieren, werden zum einen durch die Aktivität der intrinsischen und extrinsischen Muskeln des Larynx und zum anderen durch die variierende Stärke des subglottalen Luftdrucks erzeugt. Die Erhöhung des Luftdrucks und/oder der Stimmlippenspannung führen dabei zur gesteigerten Schwingungsfrequenz (für eine detaillierte Darstellung der physiologischen Zusammenhänge bei der Stimmerzeugung s. z.B. Brosnahan und Malmberg, 1970: 29ff.).

Das physikalische Resultat der melodischen Stimmproduktion ist ein quasi-periodisches Signal, welches auf der **akustischen Realisierungsebene** in Form von meßbaren Änderungen der sogenannten Grundfrequenz (F0) reflektiert wird. Das von den Stimmlippen erzeugte Signal ist allerdings nicht unmittelbar der Messung zugänglich: Es kann zum einen aus dem laryngographischen Signal ermittelt werden, welches den gemessenen elektrischen Widerstand bei unterschiedlichem Kontaktgrad bzw. der Entfernung der Stimmlippen darstellt. Zum anderen wird es aus einem komplexen Sprachsignal extrahiert, welches auch die Eigenfrequenzen des beim Sprechen ins Schwingen versetzten Vokaltraktes enthält. Die akustische Meßgröße der Grundfrequenz (in Hz) ist der Periodenkehrwert der im Signal enthaltenen Periodizität, die der Anzahl der Schwingungswiederholungen der Stimmlippen pro Sekunde entspricht. Eine langsamere Schwingungsfrequenz spiegelt sich in einer tieferen Grundfrequenz wider. Die Grundfrequenz der Sprechstimme variiert bei Frauen zwischen ca. 180 und 400 Hz, bei männlichen Sprechern dagegen zwischen ca. 60 und 240 Hz (Cruttenden,

1997: 3). Physiologisch wird die untere Begrenzung der Stimme durch die Masse der Stimmlippen determiniert, während die obere Stimmhöhe durch die Elastizität der Knorpeln und Muskeln des Larynx eingeschränkt wird (nach Helfrich, 1985: 36).

Die im akustischen Sprachsignal enthaltenen Informationen werden auf der *Perzeptionsebene* dekodiert, wobei die Grundfrequenzänderungen über der Zeit mit der Wahrnehmung von Tonhöhenvariationen durch den Hörer korrelieren. Das vom Hörer wahrgenommene Intervall zwischen zwei Tonhöhen hängt jedoch nicht vom absoluten Frequenzunterschied ab, sondern vom Verhältnis der beiden Frequenzwerte. Damit ein Ton doppelt so hoch im Vergleich zu einem anderen wahrgenommen wird, muß der Frequenzunterschied um so größer sein, je höher die absoluten Werte der zu vergleichenden Frequenzen sind. Dieses Faktum führt zu einer logarithmischen Grundfrequenzskalierung, welche in der Musiktheorie seit langem verwendet wird. Der perzeptive Abstand, der bei der Verdoppelung einer Frequenz entsteht, entspricht einer Oktave. Sie wird *per definitionem* in 12 Halbtöne untergliedert, wobei sich ein Halbton (im Folgenden HT) als  $1 : \sqrt[12]{2}$  zur Grundfrequenz verhält (vgl. Sundberg, 1979). Für die Erforschung der sprachlichen Zusammenhänge ist diese Erscheinung insoweit von entscheidender Bedeutung, als die Verwendung des Halbtons perzeptive Distanzen zwischen unterschiedlichen Frequenzen unabhängig von deren absoluten Werten in Hz vergleichbar macht. So können dann auch die Stimmen verschiedener, männlicher und weiblicher Sprecher direkt gegenübergestellt werden.

### **2.1.3. Physiologische Bedingungen**

Aufgrund einer Reihe experimenteller Befunde zu den Leistungen des menschlichen Gehörs unter Laborbedingungen kann die Schlußfolgerung gezogen werden, daß das perzeptive Diskriminierungsvermögen tonaler Erscheinungen deren Kontrollierbarkeit in der Produktion übertrifft (Helfrich, 1985: 89). Die durch den Sprecher beabsichtigten tonalen Veränderungen können nicht beliebig fein kontrolliert werden. Ihre produzierbare Geschwindigkeit variiert in Abhängigkeit von Geschlecht und Training des Sprechers sowie von Richtung und Intervall der Tonbewegung, wobei einerseits für weibliche, andererseits für trainierte Personen schnellere Ausführungen der intendierten Tonvariationen registriert wurden (Sundberg, 1979). Die fallenden Tonbewegungen werden generell schneller produziert als die steigenden: Der Unterschied zwischen deren maximalen Geschwindigkeiten verhält sich bei der Intervallgröße von einem Halbton als 60 HT/s zu 40 HT/s (Leonard und Ringel, 1979). Bei

der Erhöhung des beabsichtigten Intervalls auf 6 HT steigen die maximalen Produktionsgeschwindigkeiten bis zu 120 HT/s für Abstiegsbewegungen bzw. 65 HT/s für Anstiege (Leonard und Ringel, 1979). Bei den Veränderungsintervallen zwischen 2 und 12 Halbtönen kann eine Produktionsgenauigkeit von ca. 0,5 HT erreicht werden (Sundberg, 1979). Allgemein wird der Stimmumfang eines erwachsenen Menschen auf ca. zwei Oktaven, bei geübten Sängern auf drei Oktaven beschränkt (nach Helfrich, 1985: 36).

Die Frequenzgrenzen des tonalen Hörbereichs eines erwachsenen Menschen liegen zwischen 20 Hz und 15000 Hz (Handel, 1989: 65). Ein Schall mit einer Wiederholungsfrequenz unter 20 Hz wird nicht tonal, sondern als eine Reihe von einzelnen Aufschlägen wahrgenommen. Ein Signal von mehr als 15000 Hz kann das menschliche Ohr im Durchschnitt nicht hören. Psychoakustische Untersuchungen mit synthetisierten Vokalen, deren Grundfrequenz sich im Bereich der menschlichen Sprechstimme (s. oben) bewegt, zeigen, daß die Hörempfindlichkeit für minimale Frequenzunterschiede (*just-noticeable difference, JND*) sehr fein ist (Klatt, 1973): Das Auflösungsvermögen für *steady-state*-Vokale mit ebenem F0 liegt bereits bei 0,3 Hz. Die Hörleistung verschlechtert sich geringfügig im Zusammenhang mit diphthongalen Veränderungen im Vokalspektrum (JND bei 0,5 Hz). Die Wahrnehmung eines Tonhöhenunterschieds bei einer gleitenden F0-Bewegung über dem Vokal wird schlechter, wobei die JND annähernd proportional zum Neigungsgrad der Bewegung steigt: Bei einer flachen Neigung (12 Hz/s) beträgt sie 1,5 Hz, bei einer steilen Bewegungsneigung (32 Hz/s) dagegen 4,0 Hz (Klatt, 1973). Die Richtung der zu bewertenden F0-Bewegung läßt sich hierbei jedoch leicht wahrnehmen. Das auditive Auflösungsvermögen läßt ebenfalls nach, wenn statt Niveauunterschieden Bewegungsabweichungen bewertet werden müssen ('t Hart, 1981): Die JND beträgt ca. 1,5 bis 3 HT, wobei die Unterschiede in fallenden Bewegungen schwieriger zu bewerten sind als die in Anstiegen. Nach 't Hart (1981) sind tonale Differenzen von weniger als 3 HT als irrelevant für die sprachliche Kommunikation zu betrachten.

Nicht alle im Sprachsignal enthaltenen F0-Schwankungen werden vom Sprecher beabsichtigt produziert. Einige lokale Grundfrequenzfluktuationen werden durch eine Kombination von physiologischen Faktoren verursacht und sind weitgehend sprachunabhängig: Dazu gehören sowohl vokalspezifische (sogenannte intrinsische), als auch koartikulationsbedingte F0-Unterschiede im natürlichsprachlichen Signal, die unter dem Begriff *Mikroprosodie* zusammengefaßt werden. So zeigen hohe, geschlossene Vokale in Experimenten ein konsequent höheres F0 im Vergleich zu tiefen, offenen Vokalen (z.B. Ladd und Silverman, 1984), was als Folge der unterschiedlichen Zungen- und Kieferstellung bei den

Vokalartikulationen erklärt wird (vgl. Whalen und Levitt, 1995). Die Ausprägung des Unterschieds zwischen hohen und tiefen Vokalen variiert in Abhängigkeit von mehreren Faktoren (Geschlecht des Sprechers, segmentelle Umgebung, Akzentuierung) zwischen 0,69 und 2,33 HT (Ladd und Silverman, 1984). Auch die Stimmhaftigkeit der prävokalischen Obstruenten wirkt sich auf den F0-Einsatz des nachfolgenden Vokals aus, wobei nach stimmhaften Obstruenten eine tiefere Grundfrequenz vorliegt als nach stimmlosen (s. z.B. Löqvist et al., 1989). Die beobachtete Differenz wird auf eine höhere Spannung entlang der Stimmlippen während der Produktion stimmloser Konsonanten zurückgeführt (Löqvist et al., 1989). Für Plosive werden zusätzlich aerodynamische Effekte während der Verschlußbildung und -lösung (speziell der erhöhte Luftstrom bei der Lösung stimmloser Plosive) als mögliche weitere Gründe für die registrierten F0-Unterschiede angeführt (vgl. Hombert et al., 1979).

Das menschliche Gehör scheint diese minimalen intrinsischen und koartikulatorischen F0-Unterschiede wahrnehmen zu können (vgl. oben), wobei sie im Prozeß der Sprachverarbeitung jedoch eine andere Rolle spielen als die Wahrnehmung globaler, sprecherintendierter F0-Veränderungen auf der Äußerungsebene (auch *Makroprosodie*). Die mikroprosodischen Informationen werden bei segmenteller Dekodierung eines ambigen Sprachsignals relevant (Reinholt Petersen, 1986), während sie auf der makroprosodischen Ebene eher kompensiert werden (Gartenberg und Panzlaff-Reuter, 1991). Dies führte zur Annahme einer hoch-gradigen Flexibilität des perzeptiven Sprachsystems, welches in Abhängigkeit von den situativen Erfordernissen zwischen den beiden Ebenen der Signaldekodierung umschalten kann (Reinholt Petersen, 1986).

#### **2.1.4. Intonationskategorien**

Im Zentrum der vorliegenden Arbeit soll der makroprosodische Sprachkode stehen. Die Intonation ist in diesem Sinne eine tonale bzw. melodische Variation mit segmentübergreifender zeitlicher Ausdehnung, die aus einem kontrollierten Einsetzen des Sprechapparates resultiert und in Rahmen verschiedener kommunikativer Situationen zum Ausdruck und/oder zur Verdeutlichung der Sprecherabsichten dient. Die zu unterscheidenden, kommunikativ relevanten Intonationseinheiten einer Sprache beschrieb Pike (1945: 53, Hervorhebung im Original):

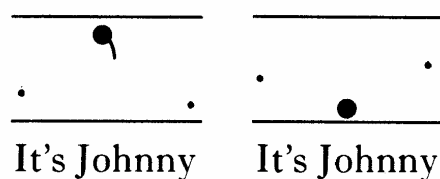
*„In each language ... the use of pitch fluctuation tends to become semi-standardized, or formalised, so that all speakers of the language use basic pitch sequences in similar ways under similar circumstances. These abstracted characteristic sentence melodies may be called **intonational contours**“.*

Theoretische Abhandlungen (z.B. Halliday, 1967) sowie empirische Untersuchungen (Isačenko und Schädlich, 1970; 't Hart et al., 1990; Kohler, 1991a) zu sogenannten *Akzentsprachen* (engl. *stress languages*, vgl. Hirst und Di Cristo, 1998: 9), zu denen auch das Russische zählt, zeigen auf, daß sich die systematische intonatorische Musterbildung in erster Linie an den Akzentstellen verankert und in Relation zum Akzentnukleus (Vokal bzw. Silbe) betrachtet werden sollte. Thorsen (1979: 72) definiert Intonationskonturen explizit als „*the gross course described by the stressed syllable alone*“. Obwohl die Akzentuierung und die Intonationseinheiten in vielen Sprachen häufig durch die gleichen Formmittel zum Ausdruck gebracht werden, sind sie als getrennte Phänomene zu betrachten. Die auditive Wahrnehmung der Prominenz einer Silbe in Relation zum Rest der Äußerung, die in das linguistische Konzept des Akzents als einer funktional relevanten Hervorhebung in der Informationsstruktur einer Äußerung (vgl. Bolinger, 1972: 633) mündet, kann auch durch nicht-tonale Mittel hervorgerufen werden: beispielsweise durch Erhöhung der relativen segmentellen Vokaldauer und/oder -intensität (vgl. Lehiste, 1970: 125; Svetozarova, 1982: 114f.) oder durch Steigerung des artikulatorischen Aufwands auf der Akzentsilbe (sogenannter *Druckakzent*, engl. *force accent*; Kohler, 2003). In einigen Sprachen (wie z.B. dem Russischen und dem Englischen) manifestiert sich der Akzent unter anderem durch die Vokalqualität, wobei eine Vokalreduktion in unakzentuierten Silben zu vermerken ist (Lehiste, 1970: 139ff., Bondarko, 1998: 220ff.).

In der modernen Intonationsforschung lassen sich zwei Beschreibungstraditionen im Hinblick auf die von der phonetischen Vielfalt abstrahierten intonatorischen Kategorien unterscheiden: Im Rahmen des holistischen Ansatzes werden die ***Tonhöhenbewegungen*** (*steigend, fallend, eben*) als sprachlich relevante Einheiten aufgefaßt, welche komplexere tonale Muster (wie z.B. *Gipfel* und *Täler*) konstituieren (z.B. Halliday, 1967; Kohler, 1991a; Kodzasov, 1996). Der kompositionelle Ansatz geht dagegen von einer begrenzten Anzahl von ***Tonhöhenebenen*** (z.B. *extra-high, high, mid* vs. *low* bzw. nur *high* vs. *low*) aus, zu denen jede Satzmelodie dekomponiert werden kann (vgl. Pike, 1945; Pierrehumbert, 1980; Féry, 1993). Die beiden Traditionen unterscheiden sich nicht nur auf der reinen Beschreibungsebene, sondern auch in der Betrachtung der funktionalen Ebene von abstrakten, sprachabhängigen Intonationseinheiten: Gemäß dem Tonstufen-Ansatz interagiert die Intonationsstruktur einer Äußerung weitgehend mit der Syntax und ist ableitbar (vgl. Pierrehumbert, 1980; Féry, 1993). Durch die Vertreter des holistischen Ansatzes wird dagegen eine Beziehung zwischen intonatorischen Formen und deren semantisch-pragmatischen Einbettung angestrebt (vgl. Kohler, 1991a, Kodzasov, 1996). In der vorliegenden Arbeit wird die letztere

Sichtweise der Intonation vertreten, die sich in der oben erläuterten, zentralen Rolle der Intonation für die lautsprachliche Kommunikation begründet. Einen weiteren Punkt, der gegen die Annahme des dekomponierenden Tonstufen-Prinzips im kommunikativen Rahmen der vorliegenden Arbeit spricht, bildet die in der Tradition von Pierrehumbert (1980) vertretene Auffassung, daß die hohen und tiefen *Töne* sowie ihre Zusammensetzungen lokal fungieren und keiner Planung auf globaler (Äußerungs-)Ebene unterliegen. Es gibt jedoch eine Reihe von Hinweisen aus empirischer, phonetisch sowie psycholinguistisch basierter Forschung, daß ein Sprecher den Intonationsverlauf von Äußerungen vorausplant (vgl. Gårding, 1983; Thorsen, 1985; Zimmer, 1988). Derartige theoretische Überlegungen führen auch in dieser Arbeit zu der Auffassung, daß „*speech pitch can best be described in terms of pitch movements rather than pitch levels*“ (‘t Hart, 1981: 820).

Abgesehen von den mit Akzentnuklei synchronisierten tonalen Mustern wird in beiden Ansätzen der Tonhöhenverlauf äußerungsinital vor dem ersten Akzentmuster (Vorlauf) sowie



die finale Tonhöhenlage nach dem letzten Akzent (Nachlauf) als bedeutsam für die adäquate Beschreibung von Äußerungsintonation betrachtet. Die Tonhöhenverläufe am Äußerungsanfang und –ende sowie über den Akzentnuklei können also das System der distinktiven, kommunikativ relevanten Intonationseinheiten einer Sprache konstituieren. Das Repertoire dieser abstrakten (phonologischen) Melodietypen ist sprachabhängig. Die Abbildung 2./2. stellt z.B. ein Schema zweier Intonationsmuster (Gipfel- und Talkontur) dar, die in mehreren Akzentsprachen beobachtet wurden (vgl. Peters, 1998; Rossi, 1998).

Abb. 2./2.: Schematische Darstellung eines steigend-fallenden Intonationsmusters mit dem Tonhöhenmaximum in der Akzentsilbe (*Gipfelkontur*, links) vs. eines fallend-steigenden Musters mit dem Tonhöhenminimum in der Akzentsilbe (*Talkontur*, rechts) über der englischen Äußerung „*It's Johnny*“ (nach Cruttenden, 1997: 10).

Phonologisch gleiche Melodietypen verschiedener Sprachen werden auf der akustischen Ebene meist durch unterschiedliche F<sub>0</sub>-Parameterausprägungen charakterisiert. Die akustischen Merkmale einer F<sub>0</sub>-Kontur, welche die Funktion der melodischen Distinktivität im Intonationssystem einer Sprache erfüllen können, faßt z.B. Torsujeva (1979: 66f.) wie

folgt zusammen: (1) Gestalt der F0-Änderungen; (2) Lage der F0-Änderungen; (3) Lage des Einsatzes und Endes einer F0-Kontur; (4) Intervall zwischen zwei Konturpunkten; (5) Umfang der F0-Änderungen; (6) Geschwindigkeit der F0-Änderungen; (7) F0-Maxima. Die perzeptive sowie kommunikative Relevanz kommt hierbei nicht den absoluten F0-Größen, sondern den relativen Werten der einzelnen Grundfrequenzparameter über einer Äußerung zu.

Die Funktion der intonatorischen Eigenschaften variiert in den Sprachen der Welt sehr stark. So zeigt beispielsweise eine komparative Studie von Nikolajeva (1989), daß in finno-ugrischen Sprachen drei unterschiedliche Intonationskonturen (steigend-fallend vs. fallend-steigend vs. steil steigend) die Funktion erfüllen können, verschiedene Fragetypen zu markieren. Die funktionale Zuordnung einer intonatorischen Kontur zu einem bestimmten Fragetyp ist hierbei sprachabhängig (Nikolajeva, 1989: 13). Die Universalienforschung im Bereich der Intonation erscheint vielen Linguisten und Phonetikern sehr attraktiv, erzielte bis jetzt jedoch keine weitreichenden Ergebnisse, was u.a. auch auf grundlegende theoretische Unterschiede innerhalb der einzelsprachlichen Intonationsbeschreibungen zurückgeführt werden kann (vgl. Ladd, 1996: 113).

### **2.1.5. Methoden der Intonationsforschung**

So vielfältig das Problemfeld der Intonation ist, unterscheiden sich auch die Forschungsansätze und deren praktische Methoden der Datenerhebung und -auswertung. Die historische Entwicklung der Intonationsanalyse erstreckt sich von impressionistischen Beschreibungen von Melodiekonturen in Einzelsprachen (z.B. Bolinger, 1978; von Essen, 1964) zunehmend zur instrumentell gestützten Experimentalphonetik.

Auf der artikulatorischen Manifestationsebene erfassen apparative physiologische Untersuchungen das Kehlkopf- und Atmungsverhalten beim Sprechen: Die Messung der Muskelaktivitäten im Larynx, ermöglicht durch neuere Entwicklung der Elektromyographie (vgl. Hast, 1975), sowie die Messungen der Luftströmung, des subglottalen und intraoralen Drucks durch Anwendung von Pneumotachographie und Manometrie können bei Klärung der kontrollierten, kommunikativ relevanten im Gegensatz zu physiologisch determinierten Erscheinungen bei der Tonproduktion helfen (vgl. Ohala, 1978). Diese Methoden haben eine Reihe von Nachteilen: Da sie (z.B. im Fall von Elektromyographie oder Laryngoskopie) mit

dem Eindringen in den Körper der Probanden verbunden sind, sinkt die Anzahl der letzteren im Vergleich zu nichtinvasiven Verfahren drastisch. Zum anderen sind die Möglichkeiten der Produktion von Spontansprache unter den Bedingungen eines auf invasiven Methoden basierenden Experiments sehr eingeschränkt, so daß die Beobachtung der Intonationsphänomene aus der Alltagssprache ausgeschlossen bleibt. Die erwähnten Nachteile veranlassen viele Autoren dazu, die Erkenntnisgewinnung auf der akustischen und perzeptiven Ebene anzusetzen und eine hypothesengeleitete physiologische Datenerhebung als die letzte Stufe der intonatorischen Theoriebildung anzusehen (s. z.B. 't Hart et al., 1990: 39).

Die meisten neueren Untersuchungen zur intonatorischen Realisierung verschiedener Kommunikationsabsichten durch den Sprecher erfolgen mithilfe der akustischen Analyse von Sprachsignalen. Die modernen Methoden der digitalen Signalverarbeitung ermöglichen beispielsweise eine rechnergestützte Bestimmung der physikalischen Korrelate der Intonation (Grundfrequenz, s. 2.1.2.), welche die rein subjektive impressionistische, vielfach kritisierte Vorgehensweise der Melodiebeschreibung ergänzen (vgl. Liberman, 1965; Adriaens, 1991: 12f; Takefuta, 1975). Da ein natürlichsprachliches akustisches Signal jedoch sehr komplex ist, müssen für die Gewinnung der abstrakten Intonationseinheiten aus der beobachtbaren Vielfalt der phonetischen F<sub>0</sub>-Schwankungen eindeutige Prinzipien der Musterextraktion bestimmt werden: Dies geschieht (1) durch eine Beschränkung des Untersuchungsmaterials auf (konstruierte) durchgehend stimmhafte Sprachbeispiele (Makarova, 1999), (2) durch eine gezielte Eliminierung der Mikroprosodie (Thorsen, 1979; 't Hart et al., 1990) oder (3) durch eine bewußte Berücksichtigung der mikroprosodischen Einflüsse (Kohler, 1991b).

Eine adäquate Beschreibung der intonatorischen Kategorien einer Sprache soll im Hinblick auf die kommunikative Relevanz der Intonation jedoch nicht nur die Produktion der Melodiemuster durch den Sprecher, sondern auch deren Perzeption und Interpretation durch den Hörer berücksichtigen. Auf der perzeptiven Ebene können psycholinguistische Tests Einblicke in die Hörerreaktionen auf das sprachliche Signal liefern. Diese Art von Tests kann die Relevanz intonatorischer Parameterausprägungen verifizieren und die akustischen Unterschiede feststellen, welche die perzeptiven Differenzierungen unterschiedlicher Sprecherintentionen bewirken. Hierbei wird eine sorgfältige methodische Planung des Perzeptions-experiments und seines Designs sowie eine anschließende statistische Bearbeitung der Daten notwendig, um zuverlässige Ergebnisse liefern zu können, denn „*a perceptual approach is subjective per se*“ ('t Hart et al., 1990: 6). Eine der in phonetischer Perzeptionsforschung etablierten Methoden stellt das Paradigma der kategorialen Wahrnehmung dar, welche traditionell als eine Korrelation zwischen Diskriminations- und Identifikationsverhalten des

Hörers definiert ist (Repp, 1984; s. McQueen (1996) für ein Überblick). Der ursprünglich in der segmentellen Phonetik verwurzelte Grundgedanke dieser Methode ist, daß ein physikalisches Kontinuum eines sprachlich relevanten akustischen Parameters nicht als eine kontinuierliche Veränderungsreihe wahrgenommen wird, sondern in Abschnitte diskreter Sprachkategorien unterteilt erscheint. Hierbei zeigt die Identifikationsfunktion eine abrupte Änderung an einer Stelle entlang des physikalischen Kontinuums, die der vorliegenden Kategoriengrenze entspricht. Parallel dazu demonstriert die Diskriminationsfunktion an der gleichen Stelle ein Maximum des Unterscheidungsvermögens der Hörer, die nur diejenigen zu vergleichenden Stimuluspaare als unterschiedlich beurteilen, deren Mitglieder bei der Identifikation zwei unterschiedlichen Kategorien zugeordnet werden (Repp, 1984: 266; s. Abb. 2./3.). Nach Repp (1984) liegt der Nachteil des Paradigmas der kategorialen Wahrnehmung darin, daß diese Methode resynthetisiertes Material verlangt und die Übertragung der Ergebnisse auf natürliche Sprache eventuell als eingeschränkt zu betrachten ist. Der Einwand der Unnatürlichkeit ist für die Experimente mit segmentellen Kontinua angebracht, die den Probanden unter lebensfernen Bedingungen präsentiert werden. Dies gilt jedoch nicht im Fall eines intonatorischen Perzeptionsexperiments mit der Beurteilung von Äußerungen auf deren kommunikative Absichten hin (vgl. Schiefer und Batliner, 1988). Das Konzept der kategorialen Wahrnehmung wird der durchzuführenden Perzeptionsuntersuchung zugrunde gelegt (Kapitel 5).

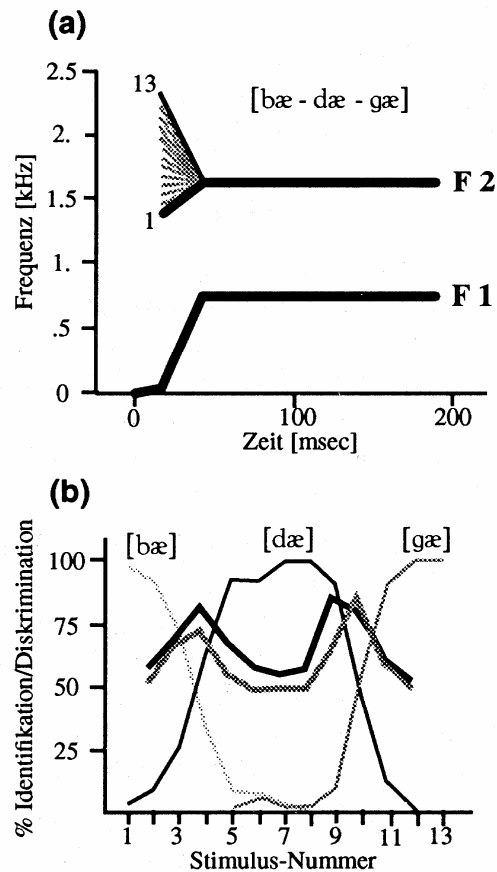


Abb. 2./3.: Der Experimentaufbau im Paradigma der kategorialen Wahrnehmung im segmentellen Bereich (nach Pompino-Marschall, 1995: 154): (a) ein Vokalkontinuum der äquidistant veränderten Startfrequenz des zweiten Formanten (F2) mit 13 Stimuli; (b) Identifikationsfunktion von [bæ]-, [dæ]-, [gæ]-Antworten sowie hieraus errechnete Diskrimination (fett grau) und die experimentell gewonnene Diskriminationsfunktion (fett schwarz).

## 2.2. Sprachliche Kategorien der Frage vs. Aussage

### 2.2.1. Semantik und Pragmatik

Für die sprachliche Kommunikation haben Fragen und Aussagen primäre Relevanz als sprachliche Handlungstypen, die einen gezielten Informationsaustausch ermöglichen. Während bei einem deklarativen Sprechakt Sachverhalte oder Meinungen mitgeteilt werden, fordert ein interrogativer Sprechakt Äußerungen und Stellungnahmen zu Sachverhalten. Durch eine Frage bringt ein Sprecher zum Ausdruck, daß er eine bestimmte Information braucht. Wunderlich (1976: 169) stellt heraus: „Der Ursprung der Frage liegt in einem

*kognitiven Defizit*“. Eine Frageformulierung äußert die Annahme des Sprechers, daß der Hörer über die benötigte Information verfügt, und stellt eine Bitte um Auskunft dar. Eine Frage involviert im hohen Maße den Hörer in den Kommunikationsprozeß, da sie eine sprachliche Reaktion in Form einer Antwort verlangt.

Die kognitive Kategorie der Frage kann in verschiedene situative Kontexte eingebettet werden, was eine Typologie der Frage notwendig macht. Im allgemeinen werden 2 grundlegende Fragetypen unterschieden: „*Die Entscheidungs- und die Ergänzungsfragesätze bilden ... das Fundamentum des erotetischen Systems...*“ (Luukko-Vinchenzo, 1988: 8; vgl. auch Cruttenden 1997: 155; v. Essen, 1964: 44). Bei einer Ergänzungsfrage (auch *partielle Frage*, *W-Frage* oder *Fragewort-Frage* genannt) handelt es sich aus semantisch-pragmatischer Sicht um eine Forderung nach einer speziellen (Teil-)Information, nach einer Präzisierung eines bereits bekannten Sachverhalts. Diese Art von Fragen zeigt eine „Wissenslücke“, die vom Adressaten ausgefüllt werden muß. Mit einer Entscheidungsfrage (auch *allgemeine Frage*, *Satz-Fragesatz*, *Ja/Nein-Frage*, *Frage ohne Fragewort*) dagegen äußert der Sprecher bereits eine Vermutung, die vom Gesprächspartner entweder bestätigt oder verneint werden soll. Die Beantwortungsmöglichkeiten der Entscheidungsfragen erschöpfen sich jedoch nicht in den binären *Ja/Nein*-Antworten, sondern erlauben Graduierungen (wie *sicher*, *vielleicht*, *wahrscheinlich*, *kaum*). Daher scheint der weit verbreitete Terminus *Ja/Nein-Frage* für diesen Fragetyp zu eng gefaßt. Eine Entscheidungsfrage soll den Hörer dazu veranlassen, den Realitätsbezug des Geäußerten einzuschätzen.

### **2.2.2. Formale Kodierung der Frage im Sprachvergleich**

Damit der Perzipient die kommunikative Intention des Sprechers adäquat interpretieren und auf sie reagieren kann, müssen sprachliche Handlungstypen wie Frage oder Aussage entsprechend gekennzeichnet werden. In der Linguistik und Phonetik werden die formalen Kenntnissysteme diskutiert, „*die dem Ausführen und Verstehen von Äußerungen in bestimmten kommunikativen Kontexten zugrunde liegen*“ (Motsch und Viehweger, 1981: 127). Dabei wird eine Übereinstimmung zwischen intendierten sprachlichen Handlungen und deren interpretierter kommunikativer Funktion angenommen. Bei der Untersuchung der Organisation einer Sprache wird der systematische Zusammenhang zwischen pragmatischen Funktionen bzw. Handlungstypen und ihren formalen Eigenschaften als *Satzmodus* bezeichnet (vgl. Altmann, 1987: 22).

Jede Sprache verfügt über mehrere formale Ausdrucksmittel zur Kodierung von kommunikativ relevanten Funktionen. So wird der Handlungsakt der Ergänzungsfrage meist lexikalisch durch ein Fragewort gekennzeichnet, welches sich auf die fehlende Teilinformation, nach der gesucht wird, bezieht. Der formale Ausdruck der Entscheidungsfrage ist dagegen in den Sprachen der Welt sehr variabel. Die Markierung kann morphologisch sein, wie z. B. eine Verbalflexion im Grönländischen. Auch durch syntaktische Mittel kann dieser Typ von Frage ausgedrückt werden. Im Deutschen oder Englischen etwa geschieht dies durch Subjekt-Verb-Inversion. Einige Sprachen setzen für die Markierung von Entscheidungsfragen spezielle Lexika ein: Das Chinesische und Karenni verfügen beispielsweise über eine satzfinale Fragepartikel, während im Kurdischen diese Funktion durch Modalwörter erfüllt wird. Es gibt auch eine Reihe von Sprachen, welche diesen Fragetyp ausschließlich durch eine Kontrastbildung auf der Ebene der Intonation von anderen sprachlichen Handlungsakten differenzieren. Die intonatorischen Kontraste können hierbei sehr unterschiedliche Ausprägungen annehmen.

In der überwiegenden Anzahl der untersuchten Sprachen wurde beobachtet, daß die Intonation in Entscheidungsfragen final steigend, in Aussageäußerungen dagegen final fallend ist (vgl. Hermann, 1942; Bolinger, 1978). Dieses Phänomen wurde oft als eine Sprachuniversalie generalisiert (vgl. *the Strong Universalist Hypothesis* von Ladd, 1981). Für diese Universalie finden sich jedoch verschieden geartete Gegenbeispiele, wie in den Satzmodussystemen des Ungarischen und des australischen Englisch. Im Ungarischen haben sowohl Aussagen als auch Fragen eine final fallende Intonation (s. Gósy und Terken, 1994), während in der australischen Varietät des Englischen ein Phänomen des sogenannten *Uptalk* beobachtet wurde (Fletcher und Harrington, 2001), bei dem Aussagen wie Fragen eine stark ansteigende Kontur gegen Ende der Äußerung aufweisen. In beiden Sprachen werden die diskutierten Satzmodi nicht allein durch die äßerungsfinale Kontur, sondern durch eine spezifische Ausprägung des Intonationsmusters über der Akzentsilbe markiert. Im australischen Englisch werden die deklarativen Äußerungen durch einen tiefen Ausgangspunkt der Anstiegsbewegung auf der Akzentsilbe gekennzeichnet. Im Gegensatz dazu wird das intonatorische Muster der Entscheidungsfragen durch einen höher einsetzenden F<sub>0</sub>-Anstieg über der Akzentsilbe geprägt. Im Ungarischen liegt der Unterschied zwischen den Intonationskonturen von isoliert produzierten Entscheidungsfragen und Aussagen in der Synchronisation der steigend-fallenden Gipfelbewegung mit den Segmenten der Äußerung (s. Gósy und Terken, 1994: 270f.). In Aussagen tritt der Gipfelhochpunkt innerhalb des Akzentvokals auf. Die Kontur fällt bis zum Äußerungsende. In Fragen liegt der

Intonationsgipfel, meist unabhängig von der Silbenzahl der Äußerung, auf der Pänultima, wobei der Gipfelabstieg sich über die Endsilbe erstreckt. Fällt der Satzakkzent auf die letzte bzw. vorletzte Äußerungssilbe, wird der Gipfel mit einem geringen Abstieg auf der letzten bzw. der einzigen Silbe realisiert. Die Kontur über dem Akzentvokal verläuft entweder auf einem tiefen Niveau oder steigt ausgehend von einem Tiefpunkt an.

In Sprachen wie dem Deutschen oder dem Englischen, die eine formal-grammatische Markierung für Entscheidungsfragen besitzen, können in Abhängigkeit von der Kommunikationssituation sowohl steigende als auch fallende Konturen auftreten. Die häufige Behauptung, Entscheidungsfragen hätten im Englischen typischerweise eine final steigende Intonationskontur, widerlegt Fries (1964), der die kommunikative Situation einer Fernsehquizsendung analysiert und dabei in der Mehrzahl der produzierten Entscheidungsfragen (62%) fallende Intonation feststellt. Eine pragmatische Interpretation des Unterschieds im Gebrauch der beiden Konturen gibt Crystal (1969: 3f.), der einen Zusammenhang zwischen dem vom Sprecher gewählten Intonationsmuster und der Art seiner persönlichen Einstellung zum Hörer sieht. Die final steigenden Konturen in Entscheidungsfragen klingen laut Crystal „freundlicher“, „interessierter“, während fallende Konturen auf den Adressaten „seriöser“ und „schroffer“ wirken sollen.

In ähnlicher Weise findet die Behauptung von O. v. Essen (1964: 15), die Entscheidungsfrage im Deutschen könne ausschließlich durch eine final steigende Kontur („*interrogative Intonation*“) markiert werden, in der modernen Intonationsforschung keine Unterstützung. Daß dies unter dem semantisch-pragmatischen Aspekt keineswegs der Fall ist, zeigt z. B. Kohler (1995). Er interpretiert die syntaktisch markierten Entscheidungsfragen mit steigender Intonation als „*interessierte Bitten um Auskunftsalternativen*“, die mit fallender Endkontur produzierten Fragen dagegen als „*Vorwegnahme der Entscheidung für eine der Alternativen durch den Sprecher*“ (Kohler, 1995: 197).

In Sprachen mit grammatischen Satzmodus-Markern können auch Sonderfälle wie syntaktische Ellipsen oder Deklarativsätze in der Funktion von Entscheidungsfragen vorkommen (vgl. dt. *Für Susanne.* vs. *Für Susanne?* und *Er kommt heute.* vs. *Er kommt heute?*, s. Paul, 1898: 122; Altmann, 1987: 42f.). In einer komparativen Perzeptionsuntersuchung zum Schwedischen und zur amerikanischen Varietät des Englischen zeigen Hadding-Koch und Studdert-Kennedy (1964), daß die Interpretation einer isoliert präsentierten Äußerung mit der syntaktischen Struktur einer Ellipse als Frage vs. Aussage in beiden Sprachen vollkommen von der Intonation abhängt. Bei den getesteten Gipfelkonturen, die anschließend eine einfach fallende oder eine fallend-steigende Bewegung aufweisen, ist nicht

nur die äusserungsfinale F0-Kontur von entscheidender Bedeutung, sondern auch die Höhe des vorangehenden Gipfels und der Wendepunkt der fallend-steigenden Grundfrequenzbewegung. Die Interaktion dieser drei Parameter scheint sich auf die Frage- vs. Aussageperzeption auszuwirken (s. Hadding-Koch und Studdert-Kennedy, 1964: 181f.). Dabei werden von schwedischen und amerikanischen Hörern differente Konturen bevorzugt als Fragen oder als Aussagen gedeutet: Amerikanische Hörer brauchen im Vergleich zu schwedischen Hörern eine höher steigende finale Kontur, um zur eindeutigen Fragewahrnehmung zu gelangen. Sie tendieren auch dazu, Äußerungen mit geringfügig ansteigenden Endkonturen als Aussagen wahrzunehmen, während schwedische Muttersprachler nur Äußerungen mit final fallenden Konturen als Aussagen akzeptieren.

Die angeführten Beispiele zeigen, daß die formal möglichen Ausdrucksmittel im Hinblick auf die Kodierung der kommunikationsrelevanten Funktionen wie Aussage und Entscheidungsfrage in den Sprachen der Welt unterschiedlichen Stellenwert besitzen. Dabei spielt der intonatorische Kode, welcher sowohl autonom als auch in der Kombination mit anderen Formmitteln fungieren kann, eine bedeutende Rolle. In den Sprachen der Welt werden unterschiedlich gestaltete, in systematischem Kontrast zueinander stehende Intonationsmuster verwendet, um Satzmodi zu markieren oder kommunikative Situationskontexte innerhalb eines grammatisch gekennzeichneten Satzmodus zu differenzieren. Dabei handelt es sich speziell um Grundfrequenzkonturen auf der Akzentsilbe und am Äußerungsende.

Besonders relevant für die linguistische und phonetische Theoriebildung ist die Intonationsforschung in den Sprachen, in denen ausschließlich Intonation zur Satzmodusdifferenzierung eingesetzt wird, wie z.B. im modernen Griechisch, Portugiesisch, Jacaltec (vgl. Cruttenden, 1997: 155). Auch das moderne Russische zählt hierzu.

Unter den Formmitteln, die zwischen Aussage und Entscheidungsfrage (im Folgenden synonym mit „Frage“) im Russischen unterscheiden können, führen moderne russische Grammatiken (Shviedova, 1970: 570f.; Rosental, 1979: 24ff.; E.-G. Kirschbaum, 2001: 309f.) in erster Linie die Intonation an. Hierbei wird oft eine „*besondere Frageintonation*“ (s. Rosental, 1979: 24) genannt, mit deren Hilfe eine Aussage in eine Entscheidungsfrage „*umgewandelt*“ werden kann, ohne daß die lexikalische oder syntaktische Form der Äußerung geändert werden. Diese „*Frageintonation*“ wird durch einen steilen Tonanstieg auf dem „*logisch hervorgehobenen Wort*“ charakterisiert. Andere Fragemarker, wie Partikel oder syntaktische Satzanfangsstellung des in der Frage fokussierten Wortes, sind auch möglich, enthalten jedoch entweder emotionale oder stilistische Konnotationen.

Die Differenzierung zwischen Aussage und Entscheidungsfrage im Russischen beruht ganz überwiegend auf der intonatorischen Gestaltung der Äußerungen, wobei der Unterschied mit den sprachuniversalen Tendenzen (Ladd, 1981) nicht konform zu gehen scheint. Da die hier zitierten russischen Grammatiken (s. oben) nur sehr ungenaue Angaben über die Gestalt der Intonationsmuster in Frageäußerungen und keine Beschreibungen der intonatorischen Formen von Aussageäußerungen liefern, erhebt sich die Frage, welcher Kontrast im Intonationssystem des Russischen entstanden ist, so daß eine Äußerung als Aussage bzw. Frage produziert und interpretiert werden kann.

### **3. Charakteristika der russischen Intonationsmuster aufgrund der akustischen Analyse von Produktionsdaten**

In diesem Kapitel werden zunächst einige Ansätze zur Beschreibung der Intonation des Russischen diskutiert sowie die Notwendigkeit einer ergänzenden akustischen Voruntersuchung begründet (3.1.). Diese wird im Unterkapitel 3.2. beschrieben.

#### **3.1. Beschreibungen der russischen Intonation in der Literatur**

Es existiert eine große Anzahl an Publikationen zur Intonation des Russischen, die sich von phonetischen Untersuchungen zu ausgewählten Problemen bis hin zur Aufstellung eines phonologischen Systems der russischen Intonation erstrecken. Die Ansätze lassen sich als auditiv-impressionistisch, phonetisch-akustisch oder phonologisch klassifizieren. Im Folgenden wird ohne Anspruch auf Vollständigkeit eine Auswahl an einschlägigen Publikationen mit Blick auf den für diese Arbeit relevanten Aspekt der formal-intonatorischen Kodierung eines Satzmodusunterschieds beleuchtet, wobei sowohl ältere als auch neuere Untersuchungen sowie etablierte Intonationsmodelle beachtet werden. Nach einer kurzen Charakterisierung eines Ansatzes wird auf die spezifische Unterscheidung zwischen Fragen und Aussagen kritisch eingegangen.

Die frühen Arbeiten zur russischen Intonation haben in erster Linie den didaktischen Charakter von Lehrbüchern. Dieser Ansatz kann als impressionistisch bezeichnet werden, da Intonationsverläufe einem Höreindruck entsprechend verbal und graphisch dargestellt werden. Unerheblich für den Ansatz ist die Unterteilung des beim Sprechen benutzten melodischen Umfangs einer Stimme in mehrere (meistens drei bis vier) tonale Ebenen, welche einen Rahmen für die klassifizierende Beschreibung aller intonatorischen Ereignisse bildet, wie z.B. in „*The Phonetics of Russian*“ von Jones und Ward (1969). Die Autoren behandeln Intonation als ein grammatisches Mittel, mit dessen Hilfe kommunikative Satztypen gestaltet werden. Der grundlegende Unterschied zwischen der melodischen Kontur einer Aussage (*plain statement*) und einer Entscheidungsfrage (*non-specific question*) liegt nach Jones and Ward (1969) im tonalen Verlauf auf der letzten betonten Silbe, die das sogenannte logische Gewicht (*logical weight*) trägt: In Aussagen wird eine fallende Tonhöhenbewegung mit dem Abstieg vom mittleren bis zum tiefen Bereich der Sprecherstimme beobachtet. In Entscheidungsfragen dagegen hat die letzte betonte Silbe eine vom mittleren Niveau bis zur oberen Grenze der Stimme steigende Kontur. Der starke Anstieg wird gefolgt durch einen

schnellen Abstieg, welcher oft am Ende der betonten Silbe beginnt. Alle nachfolgenden unbetonten Silben liegen sowohl in Aussagen als auch in Fragen an der unteren Grenze der Sprechstimme.

Eine mit Jones und Ward (1969) vergleichbare Vorgehensweise zeigt Wenk (1975) in einem Lehrbuchartikel zur Einführung in die russische Sprache. Hier wird jedoch eine höhere Abstraktionsstufe der zu klassifizierenden Intonationsverläufe angestrebt. Den Gegenstand der intonatorischen Beschreibung bilden distinktive Einheiten der *Intoneme*, welche als typisierte intonatorische Konturen über den Bereich des Hauptakzents definiert sind (Wenk, 1975: 178). Ihre Funktion gilt wie bei Jones und Ward (1969) der Gestaltung von kommunikativ definierten Äußerungstypen. Wenk unterscheidet unter anderem ein *Intonem der Aussage* mit der Funktion, Aussage und Abgeschlossenheit zu signalisieren, und ein *Intonem der Entscheidungsfrage* mit der Funktion, den Gesprächspartner zur Stellungnahme über die geäußerte Vermutung aufzufordern. Des weiteren werden für das Intonem der Aussage zwei Unterarten mit ihren speziellen formalen Ausprägungen gegeneinander abgegrenzt: *Aussage ohne besondere Hervorhebung des Rhemas*, die als Grundform dieses Intonems genannt wird, und *Aussage mit besonderer Hervorhebung des Rhemas*. Der funktionale Unterschied besteht in der kontrastierenden, verstärkten Akzentuierung. Formal werden beide Unterarten anhand der jeweiligen Tonführung über der Akzentsilbe differenziert: Charakteristisch für eine Aussage ohne Hervorhebung ist eine Tonsenkung über der Akzentsilbe, wobei auf der nachfolgenden unakzentuierten Silbe eine tiefe Tonebene erreicht wird. Eine Aussage mit Hervorhebung zeichnet sich dagegen durch eine Aufwärtsbewegung von sogenannter neutraler zu mittlerer Tonlage auf der Akzentsilbe aus, gefolgt von einer Abwärtsbewegung, die zu einer tiefen Tonebene in der nachfolgenden unakzentuierten Silbe führt.

Die formale Besonderheit des Intonems der Entscheidungsfrage liegt laut Wenk (1975: 188f.) darin, daß ein sehr großes Tonintervall zwischen der vorangehenden und der akzentuierten Silbe auftritt. Dieses Intervall wird beschrieben als ein Tonsprung zur Akzentsilbe hin und ein Tonanstieg auf dem Akzentvokal bis an die obere Grenze des Stimmumfangs. Die tonale Gestaltung des Nachlaufs kann in Abhängigkeit von der Anzahl der nachfolgenden unakzentuierten Silben unterschiedlich ausfallen: Folgen dem Akzent keine weiteren Silben, endet das Muster hoch. In sonstigen Fällen findet eine Tonsenkung mit variierender Intervallgröße statt.

Impressionistische auditive Intonationsbeschreibungen dieser Art eignen sich zweifelsohne zur Einführung in den Sprachgebrauch des Russischen im Rahmen eines Fremdsprachenunterrichts. Einer akustisch ausgerichteten Untersuchung zur intonatorischen Aus-

prägung von sprachlich relevanten Kontrasten können sie jedoch nicht zugrunde gelegt werden, da die derartigen unspezifizierten Bezüge auf Tonlage sowie Intervallgröße schwer zu interpretieren sind. Liberman (1965) kritisiert solche Ansätze wegen der Abwesenheit einer klaren physikalischen Basis für die vorgenommenen Distinktionen und der stark inkonsistenten Anwendung der aufgestellten tonalen Kategorien durch verschiedene Bearbeiter.

Eines der bekanntesten Intonationsmodelle des Russischen, das ursprünglich ebenso als Lehrbuch für die Bedürfnisse des Fremdsprachenunterrichts entstand, sich aber auch auf einige vorangegangene experimental-phonetische Untersuchungen gründete, entwickelte Bryzgunova (1977, 1980). Im Gegensatz zu den oben dargestellten Ansätzen geht Bryzgunova von einer begrenzten Anzahl distinktiver, formal definierter Einheiten aus, den sogenannten Intonationskonstruktionen (*IKs, intonatsionnyje konstruktsii*). Ihre sprachliche Funktion besteht in der Unterscheidung verschiedener Bedeutungen von Äußerungen mit gleicher lexikalischer und syntaktischer Struktur. Das gesamte Intonationssystem des Russischen umfaßt laut Bryzgunova sieben Intonationskonstruktionen, von denen IK-1 vs. IK-3 in Zusammenhang mit den Funktionen gebracht werden, Aussagen vs. Fragen zu gestalten. Der formale Unterschied zwischen IK-1 und IK-3 wird in erster Linie durch die Grundfrequenzbewegung auf dem sogenannten *Intonationszentrum* geprägt, welches als eine Silbe mit bedeutungsrelevanten F<sub>0</sub>-Änderungen definiert ist (Bryzgunova, 1977: 17, 1980: 97). IK-1, deren Funktion unter anderem in der Markierung von Aussageäußerungen besteht, zeichnet sich durch eine abrupt oder auch kontinuierlich fallende Bewegung über dem Intonationszentrum aus. Bei der IK-3, die in erster Linie zur Gestaltung von Fragen verwendet wird, liegt eine steigende Grundfrequenzbewegung im Intonationszentrum vor. Der Offset des Vokals im Intonationszentrum kann entweder ein ebenes oder ein fallendes F<sub>0</sub> aufweisen. In beiden IKs ist die F<sub>0</sub>-Lage nach dem Intonationszentrum tiefer als davor. Laut Wenk (1975: 182) entspricht dem von ihm beschriebenen Intonemtyp *Aussage mit besonderer Hervorhebung* die IK-2 in der Klassifikation von Bryzgunova. Die formale Gestaltung und die funktionale Belastung der IK-2 fallen bei Bryzgunova (1977: 26; 1980: 109ff.) jedoch anders aus, als sie für den oben genannten Intonemtyp bei Wenk (1975: 181ff.) beschrieben werden, so daß eine direkte Parallelenbildung nicht möglich erscheint: IK-2 unterscheidet sich von IK-1 lediglich durch die Verstärkung der Wortbetonung im Intonationszentrum und wird hauptsächlich zur Markierung von Ergänzungsfragen verwendet.

Die postulierte Vorgehensweise von der analysierten akustischen Form zu deren sprachlichen Funktion setzt Bryzgunova (1977, 1980) nicht konsequent um: Wie in anderen, hier zitierten Ansätzen sind Bryzgunova's Standardbeispiele für die aufgestellten Intonations-

konstruktionen an eine begrenzte Anzahl von kommunikativ relevanten Funktionen orientiert. Das Modell beruht zwar auf der akustisch-phonetischen Analyse von Sprachdaten, die Publikationen (Bryzgunova, 1977; 1980; 1984) lassen jedoch jegliche Angaben über die methodische Grundlage vermissen, auf der die Zusammenfassung phonetischer Vielfalt von Grundfrequenzverläufen zu den Intonationskonstruktionen stattfindet. Im Übrigen fehlt eine Aufstellung der dabei berücksichtigten akustischen Parameter.

Dieser Ansatz stieß auf starke Kritik vom Standpunkt der perzeptiv ausgerichteten Evaluierung der postulierten Intonationskategorien (s. Odé, 1992). Odé unterzog die Beispiele der Bryzgunova'schen Publikation von 1984 einer Perzeptionsanalyse (s. auch Unterkapitel 4.1.) und stellte fest, daß russische Muttersprachler Varianten einer IK als perzeptiv ungleich, Varianten verschiedener IKs dagegen als perzeptiv äquivalent bewerteten. Odé (1992: 281ff.) führte die Schwäche des Modells darauf zurück, daß es so bedeutsame Merkmale wie Timing, Umfang der tonalen Bewegung und Tonlage nach dem Intonationszentrum nicht berücksichtigt, welche für die perzeptive Identifikation einer Kontur relevant sind.

Einige dieser Merkmale wurden in ein weiteres Intonationsmodell des Russischen integriert, welches für das Anwendungsgebiet der Sprachsynthese entwickelt wurde (Svetozarova, 1982; 1998). Der Ansatz stellt eine Intonationsphonologie in Form der sogenannten melodischen Kommandos (*meloditshieskije komandy*) dar, die den grundlegenden kommunikativen Typen von kurzen isolierten Äußerungen zugeordnet werden. Ähnlich wie bei Wenk (1975), Jones und Ward (1969) wird eine begrenzte Anzahl kommunikativ relevanter Funktionen von Äußerungen ausgewählt, deren intonatorische Form den Ausgangspunkt für die Modellierung bildet. Die Kommandos enthalten Informationen über die Steigung der tonalen Bewegungen, die höchsten Grundfrequenzwerte und die Dauer hoher F<sub>0</sub>-Plateaus in Gipfeln. Für kurze isolierte Äußerungen wird eine melodische Grundkontur angenommen, die durch einen langsamen kontinuierlichen Abstieg der Grundfrequenz über die intonatorische Einheit vom mittleren Niveau der Sprechstimme bis zum mittel tiefen Bereich charakterisiert ist. Der Einsatz und das Ende der Basiskontur weisen einen steilen An- und Abstieg auf. In Abhängigkeit vom Äußerungstyp wird die angenommene Grundkontur durch ein entsprechendes melodisches Kommando modifiziert, welches an die Positionierung der Akzentsilbe auf der Zeitachse gebunden ist. Ein Deklarativsatz mit einem neutralen Satzakzent hat nach Svetozarova typischerweise eine Reihe von steigend-fallenden Grundfrequenzbewegungen in Verbindung mit akzentuierten Wörtern und einen tiefen F<sub>0</sub>-Abstieg auf der Akzentsilbe des finalen Wortes (Abb. 3./1.). Ein Deklarativsatz mit einem speziellen oder logischen Akzent dagegen erhält eine sogenannte

stark zentrierte Kontur über dem speziell akzentuierten Wort (Abb. 3./2.), während die umgebenden Wörter an melodischer Prominenz verlieren.

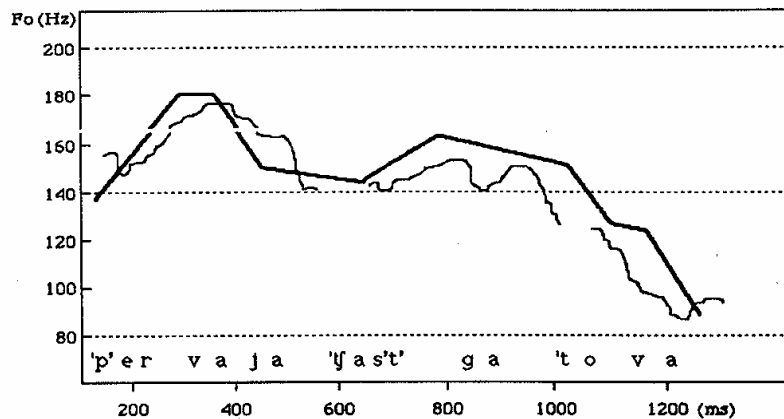


Abb. 3./1.: Die natürlich produzierte (feine Linie) vs. regelgenerierte (fette Linie) Grundfrequenzkontur der Aussageäußerung „*Piervaja tshias'tj gotova.*“ (Das **erste** Teil ist **fertig.**) (Svetozarova, 1998: 267).

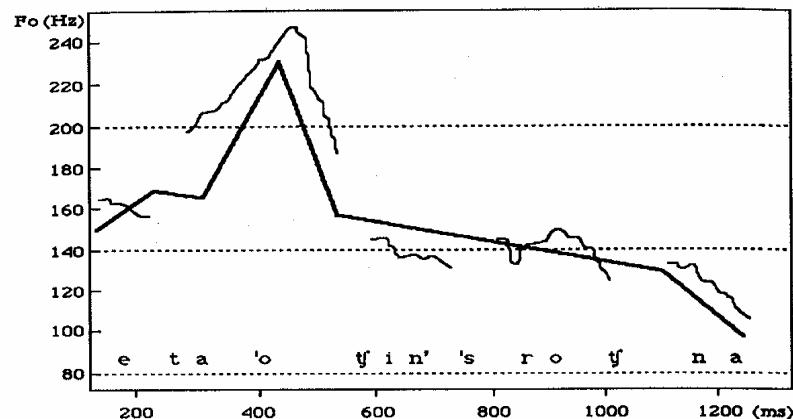


Abb. 3./2.: Die natürlich produzierte (feine Linie) vs. regelgenerierte (fette Linie) Grundfrequenzkontur der Aussageäußerung mit einem speziellen bzw. „logischen“ Akzent „*Eto otshienj srotshjno!*“ (Das ist **sehr** dringend!) (Svetozarova, 1998: 268).

Die melodische Kontur eines lexikalisch unmarkierten Interrogativsatzes schildert Svetozarova als eine vom Einsatz der Akzentsilbe des hauptakzentuierten Wortes an steil steigende Kontur, die nach einem kurzen Plateau auf dem erreichten hohen Niveau einen obligatorischen Grundfrequenzabstieg in den nachfolgenden Silben hat (s. Abb. 3./3.). Der Unterschied zwischen der Kontur eines Interrogativsatzes und der eines Deklarativsatzes mit logischem Akzent scheint ausschließlich in dem Vorhandensein eines Gipfelplateaus zu bestehen (vgl. Abb. 3./2. und 3./3.). Beide Konturen werden zwar durch unterschiedliche

Kommandos realisiert, sind jedoch sehr ähnlich beschrieben (vgl. Svetozarova, 1982: 149ff.; 1998: 267ff.). Bedauerlicherweise fehlen auch in diesen beiden Publikationen Darstellungen der Ausprägung von berücksichtigten akustischen F0-Parametern für die dargestellten Kommandoarten.

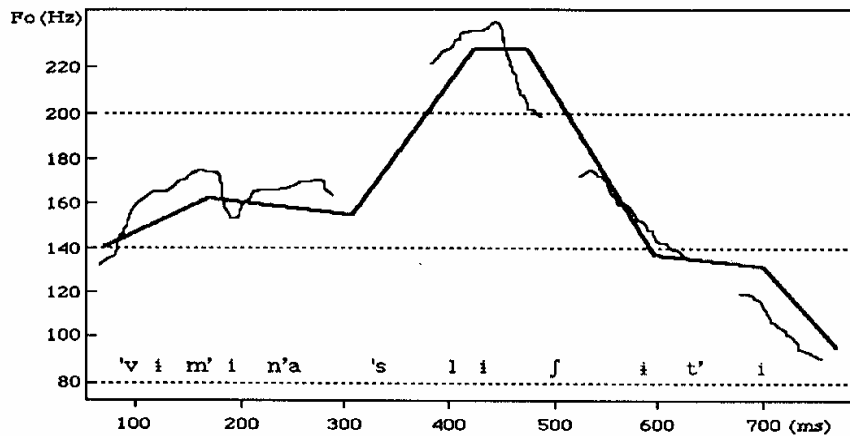


Abb. 3./3.: Die natürlich produzierte (feine Linie) vs. regelgenerierte (fette Linie) Grundfrequenzkontur der Frageäußerung „Vy mienia *slyshyite?*“ (**H**ören Sie mich?) (Svetozarova, 1998: 269).

Als einziger der hier diskutierten Ansätze strebt das Modell von Svetozarova (1982: 158ff.) eine Evaluierung der aufgestellten Muster in der Perzeption an. Die regelgenerierten Intonationskonturen wurden in Perzeptionsexperimenten auf die Identifikation der kommunikativen Typen (ausgedrückt durch die Schriftzeichen /./, /?/, /!/) und ihre Natürlichkeit hin überprüft. Bei festgestellter Interaktion der lexikalischen und intonatorischen Ebenen wurden in durchgeführten Identifikationstests besonders gute Erkennungsraten für Aussageäußerungen registriert. Weniger gute Ergebnisse lagen für Identifikation der für Entscheidungsfragen synthetisierten Konturen vor, was auf eine unpräzise temporale Lokalisierung der entsprechenden Befehle auf der Zeitachse zurückgeführt wurde. Am schlechtesten wurden die als Imperativsätze intendierten Äußerungen identifiziert. Hierbei ist unklar, welche Kontur für diesen Satztyp synthetisiert wurde, da er im Repertoire der aufgelisteten Intonationstypen nicht vorhanden ist (vgl. Svetozarova, 1982: 151). Die zu vermutende Gleichsetzung der Kontur eines Ausrufesatzes und der oben geschilderten Aussagen mit besonderer Hervorhebung wurde an keiner Stelle motiviert. Die im Experiment registrierten Urteile über die Natürlichkeit der synthetisierten Konturen korrelierten mit dem Identifikationsvermögen. Die Ergebnisse wurden dahingehend interpretiert, daß die Synchro-

nisierung einer tonalen Kontur und der segmentellen Struktur einer Äußerung von entscheidender Bedeutung ist und sich an den tontragenden Kern der Akzentsilbe orientiert. Dabei sollte die Verknüpfung des F0-Maximums einer Kontur mit dem Ende des *steady-state*-Abschnitts des Akzentvokals eine wesentliche Verbesserung der Identifikation erzielen. Ein großer Nachteil des Experiments bestand in der Anbindung der Identifikationsitems an die Interpunktion (/./, /?/, /!/), deren Norm im Russischen starke Schwankungen zuläßt (s. z.B. Savtshienko, 1962: 76f.).

Auf einer detaillierten akustischen Analyse von Einzelfällen, ausgehend von Lese-daten einer weiblichen und zweier männlicher Versuchspersonen, basiert die Publikation von Nikolajeva (1977), die unter anderem zwischen Intonation eines abgeschlossenen Deklarativsatzes (*Kadenz*) und Frageintonation (*Antikadenz*), darunter als Spezialfall Intonation der Entscheidungsfrage, unterscheidet. Für die Kadenz werden mehrere Varianten des F0-Musters innerhalb und nach der Akzentsilbe gegeben, deren Unterschiede als irrelevant aufgefaßt werden: Die Tonbewegung in der Akzentsilbe kann eben verlaufen oder ca. über 40 Hz fallen. Nach der Akzentsilbe kann das F0 auf dem gleichen, tieferen oder auch höheren Niveau im Vergleich zur Akzentsilbe liegen. Als relevant wird der F0-Unterschied im Vorlauf aufgefaßt: Der Vorlauf kann auf dem gleichen Niveau wie die Akzentsilbe eben verlaufen oder eine besondere Art von Gipfel mit dem Maximum auf der präakzentuierten Silbe bilden (*zanos*). Dieser Unterschied wird jedoch nicht weiter interpretiert. Für die Entscheidungsfrage sind ein großes tonales Anstiegsintervall von 80 bis 100 Hz in der Akzentsilbe sowie ein Abstieg der Grundfrequenz in den nachfolgenden Silben charakteristisch. Die Bewegung in der Akzentsilbe ist steigend oder steigend-eben. Bei der Analyse von Entscheidungsfragesätzen mit einem initialen Fragezentrum stellt Nikolajeva (1977: 84f.) fest, daß die F0-Lage um die letzte betonte Silbe sowohl ein terminales Kadenzmuster als auch ein nicht-terminales Muster ergeben kann.

Die Generalisierbarkeit der Beobachtungen von Nikolajeva ist insofern eingeschränkt, als sie lediglich die meßbaren Hz-Daten anführt, ohne diese unter Berücksichtigung der individuellen Sprechstimme der Versuchspersonen unterschiedlichen Geschlechts zu relativieren. Des weiteren werden die festgestellten interpersönlichen Variationen in der Realisierung der Sätze in verschiedenen Satzmodi nicht zu relevanten Mustern zusammengefaßt. Zudem werden die als relevant bezeichneten Unterschiede in keiner Weise weiter interpretiert und deren Relevanz begründet.

Eine wichtige phonetische Untersuchung zur Koordination von steigend-fallender Gipfelbewegung mit segmenteller Struktur einer Äußerung (*Timing*) führte Makarova (1999)

durch. Sie stellte fest, daß in deklarativen vs. interrogativen Ein-Wort-Äußerungen durch russische Sprecher folgende Unterscheidung im Timing einer Gipfelkontur realisiert wird: Die als Aussagen intendierten Äußerungen haben das Gipfelmaximum nah am Onset des Akzentvokals (früh), während die Frageäußerungen ihren Gipfel am Offset des Akzentvokals oder in der nachfolgenden Silbe (spät) aufweisen. Makarova (1999: 1175) weist darauf hin, daß die zeitliche Gipfelkoordination ein wichtiger Parameter für die Kodierung der Satzmodusdistinktion ist und in phonetisch sowie phonologisch begründete Beschreibungen der russischen Intonation in Produktion und auch Perzeption integriert werden sollte.

Eine andere aktuelle Arbeit zur russischen Intonation stammt von Kodzasov (1996, 1999). In diesem Forschungsansatz wird eine präzise und eingehende Explikation der sprachlich relevanten prosodischen Komponenten angestrebt, welche allerdings nicht nur tonale Merkmale umfassen. Kodzasov betont die Gewichtung verschiedener Kombinationsmöglichkeiten von mehreren prosodischen Komponenten für eine adäquate Erfassung unterschiedlicher prosodischer Muster und deren vielfältigen Bedeutungen in gesprochener Sprache. Für jede Komponentenebene wird eine funktional motivierte Parameterdifferenzierung angestrebt. Auf der Ebene der tonalen Elemente wird ein Ausgangsinventar mit drei als universal angesehenen Tonbewegungen (fallend, steigend und eben) (Kodzasov, 1996: 90) angesetzt, welche innerhalb von einer oder mehreren Silben, in einem oder über mehreren Wörtern kombiniert werden können. Dabei wird mit Bezug auf die Ergebnisse von Bryzgunova (1980) ein fallender Ton als ein rhematischer Marker in neutralen Aussagen und ein steigender Ton als Marker in Entscheidungsfragen genannt. Der Parameter des Timings wird bei Kodzasov in drei Merkmale differenziert: (1) Die F0-Bewegung erfolgt vorwiegend über einem Konsonanten und verläuft eben im Vokal (*level-like*), (2) erstreckt sich sowohl über dem Konsonant als auch über dem Vokal (*neutral*) oder (3) passiert ausschließlich über dem Vokal, der F0-Verlauf im Konsonant ist eben (*contour-like*). Als weitere relevante tonale Parameter mit ihren unterschiedlichen Ausprägungen werden Register und Intervall behandelt. Ihre sprachliche Funktion wird im expressiven Bereich angesiedelt, zur Markierung des hier zu behandelnden Satzmodusunterschieds tragen sie laut Kodzasov nicht bei. Das Kodzasov'sche kombinatorische System der prosodischen Beschreibung trägt zur Möglichkeit einer Erfassung komplexer spontansprachlicher Phänomene im Russischen entscheidend bei, die angestrebte funktionale Zuordnung der differenzierten distinktiven Einheiten beschränkt sich jedoch auf die Verallgemeinerung der Interpretation von Einzelbeispielen. In einigen Sprachbeispielen, die zur Untermauerung der postulierten Timingunterschiede angeführt werden, sind die konsonantischen Träger der relevanten F0-

Bewegung stimmlos, so daß der F0-Verlauf „rekonstruiert“ werden muß (Kodzasov, 1996: 98). Diese Differenzierung muß experimentell belegt werden.

In allen oben angeführten Untersuchungen zur russischen Intonation wird hervorgehoben, daß sich der formale intonatorische Unterschied zwischen Aussagen vs. Fragen auf der Akzentsilbe manifestiert. In den bisherigen Beschreibungen der Realisierung dieses Unterschieds wurden insgesamt drei akustische Parameter einer Grundfrequenzkontur über der Akzentsilbe berücksichtigt. Die Zusammenfassung der hierbei festgestellten charakteristischen Kontraste können tabellarisch (Tab. 3./1.) dargestellt werden.

Tab. 3./1.: Zusammenfassung der charakteristischen Unterschiede zwischen den in Fragen vs. Aussagen produzierten Grundfrequenzkonturen aus der einschlägigen Literatur.

<i>F0-Parameter</i>	<i>Aussage</i>	<i>Frage</i>
Verlauf über der Akzentsilbe	fallend	steigend
Timing	früh	spät
Intervall	klein	groß

Die Publikationen von Kodzasov (1996, 1999) und Odé (1992) weisen darauf hin, daß mit dieser Zusammenstellung möglicherweise nicht alle für die Satzmodusunterscheidung relevanten akustischen Parameter einer F0-Kontur erfaßt werden. Im Übrigen muß auch der Parameter des Intervalls präzisiert werden. Aus diesen Überlegungen scheint eine weiterführende Voruntersuchung zur Produktion von Äußerungen mit Frage- vs. Aussage-Intention notwendig.

Die Aufnahme einer speziellen Kontur für den Aussagetyp mit kontrastierender Akzentuierung ins Repertoire der meisten hier zitierten Ansätze ist bedingt durch eine in russischer Intonologie traditionell nicht vorhandene Trennung zwischen Akzent und Intonation. Sowohl Kontrastakzent als auch Satzmodus sind linguistische Kategorien, die im Russischen durch Intonation markiert werden können, stellen jedoch ganz unterschiedliche Funktionen dar. Aus diesem Grund wird in der folgenden Untersuchung auf die Behandlung einer Aussage mit Kontrastakzent verzichtet.

## **3.2. Ergänzende Voruntersuchung**

Um detaillierte Informationen zu akustischen Unterschieden in der Gipfelgestalt der als Frage vs. als Aussage produzierten Äußerungen zu erhalten, wurde die im weiteren erläuterte Produktionsvoruntersuchung geplant und durchgeführt.

### **3.2.1. Sprachmaterial und Methode**

Für die geplante Produktionsvoruntersuchung wurde der Analysebereich auf den Fall beschränkt, in dem eine Äußerung ausschließlich einen Satzakzent auf dem finalen Wort aufweist. Das Sprachmaterial wurde mit einer systematischen Variation in der segmentellen und silbischen Struktur der Äußerung konstruiert, um mögliche Auswirkungen der Segmentumgebung des Akzentvokals und der Silbenstruktur des satzakzentuierten Wortes auf die Ausprägung des zu untersuchenden Grundfrequenzmusters feststellen zu können. Die akzenttragenden Wörter enthalten mindestens zwei Silben vor der Akzentstelle (Vorlauf) und keine bis zwei Silben danach (Nachlauf). Die Akzentvokale kommen in den Äußerungen sowohl in einer Obstruenten- als auch in einer Sonorantenumgebung vor. Von insgesamt sechs ausgewählten Sätzen haben der erste und fünfte Satz eine Akzentsilbe mit einem hinteren gerundeten Vokal (/o/), in anderen Sätzen haben die Vokale eine offene Qualität (/a/). Alle Testsätze konnten sowohl als Fragen als auch als Aussagen produziert werden, wodurch sich insgesamt 12 Äußerungen ergaben.

Um den Akzentfokus und den Äußerungsmodus eindeutig zu bestimmen, wurden die Zielsätze jeweils mit einem einführenden Kontextsatz versehen, der das Thema des Zielsatzes festlegte. Das zusammengestellte Sprachmaterial (12 Satzpaare) ist samt den entsprechenden Übersetzungen ins Deutsche und der phonematisch orientierten IPA-Transkription der Zielsätze (s. Bondarko, 1998: 18f.) im Anhang B aufgelistet.

Die 12 Satzpaare wurden von einer Muttersprachlerin des Standardrussischen (Autorin dieser Arbeit) in einer randomisierten Reihenfolge zweimal vorgelesen und aufgenommen. Die Aufnahme erfolgte in der schallisolierten Kabine des Instituts für Phonetik und digitale Sprachverarbeitung Kiel. Die digitalisierten Sprachsignale wurden unter dem Sprachverarbeitungsprogramm *xassp* (URL1) bearbeitet und analysiert. Zunächst wurden die zwei Wiederholungen der 12 Zielsätze aus den jeweiligen Kontexten ausgeschnitten und in einer Signaldatei *Satzliste.wav* zusammengestellt. Anschließend wurden alle Sätze wortortho-

graphisch in lateinischer Schrift segmentiert mit Markierung des On- und Offsets des jeweiligen Akzentvokals in jedem Satz.

Die vier akustischen Parameter, durch welche die Ausprägungen der Grundfrequenzgipfel genau beschrieben werden können, wurden wie folgt gemessen: Für die Gipfelhöhe wurde der F0-Wert am Gipfelmaximum notiert, bei einem ebenen Vorlauf orientierte sich der Meßwert am letzten Punkt vor dem Grundfrequenzabstieg. Die Positionierung des Maximumpunkts in Relation zur Akzentsilbe bzw. dem Akzentvokal wurde ebenso festgehalten. Die Größe des Gipfelanstiegs ergab sich aus der Differenz zwischen dem F0-Wert am Einsatzpunkt der steigenden Bewegung und dem Maximalwert des Gipfels. Entsprechendes gilt auch für die Messungen der Abstiegsgröße. Die Steigungsgeschwindigkeit wurde aus dem produzierten Umfang der Grundfrequenzbewegung und der dafür gebrauchten Zeitspanne für An- und Abstiege des Gipfels berechnet. Bei den Messungen wurde der mikroprosodische Einfluß von Obstruenten berücksichtigt und eliminiert (s. Unterkapitel 2.1.3.). Die intrinsischen F0-Unterschiede zwischen beiden Vokalqualitäten für die Akzentvokale /a/ und /o/ (vgl. 2.1.3.) können für die aktuelle Fragestellung vernachlässigt werden, da sie bei der Berechnung von Mittelwerten in einem Korpus aufgehoben werden. Die Messungen erfolgten in Hz, anschließend wurden für alle Differenzen Halbtonwerte (HT) berechnet.

### **3.2.2. Ergebnisse**

Die tabellarisch aufbereiteten Ergebnisse der oben beschriebenen Messungen sowie Halbton-, Geschwindigkeits- und Mittelwertberechnungen sind im Anhang C zusammengefaßt.

Wie aus der Anhangstabelle C ersichtlich, sind die Grundfrequenzgipfel einer Frauenstimme in Fragen um ca. 150 Hz höher als in Aussagen, diese Differenz beträgt ca. 6 HT. Der Umfang einer Tonbewegung ist sowohl für den Anstieg als auch für den Abstieg des Gipfels in Fragen größer als in Aussagen, wobei dieser Unterschied für den Abstieg ausgeprägter ist: Diese Differenz beträgt ca. 4 HT für den Anstieg zu 6 HT für den Abstieg. Die Änderungsgeschwindigkeit in steigenden und fallenden Gipfelbewegungen ist für Frage-Äußerungen ebenfalls höher als für Aussagen, wobei die Differenz 11 HT/s im Anstieg und 13 HT/s im Abstieg ausmacht. Die fallenden Tonbewegungen verlaufen sowohl in Aussagen als auch in Fragen schneller als die steigenden. Die festgestellten Kontraste in der Gipfelgestalt können schematisch wie in der Abbildung 3./4. dargestellt werden. Die tonale Kontur in Fragen weist also eine stärker ausgeprägte Gipfelbildung als in Aussagen auf. Die

Positionierung des Gipfelmaximums auf der Zeitachse war ausnahmslos früh für die in Aussagen produzierten Gipfelkonturen und spät in Fragen.

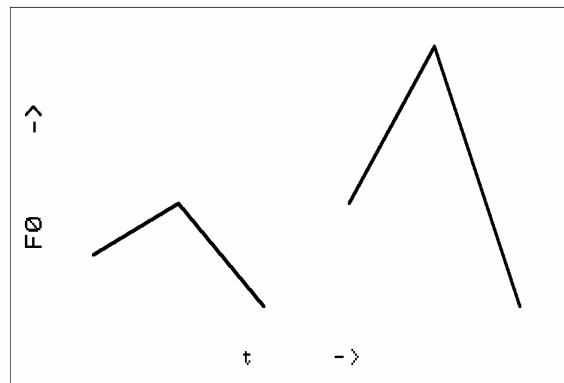


Abb. 3./4.: Ein schematischer Vergleich zwischen den im Rahmen der durchgeführten Produktionsvoruntersuchung ermittelten Gestalten von F0-Gipfeln in Aussagen (links) vs. in Fragen (rechts).

Im Zusammenhang mit der segmentellen und silbischen Struktur sind folgende Regelmäßigkeiten in Frageäußerungen aufgetreten: Bei fehlendem Nachlauf wird kein Gipfelabstieg realisiert. Falls jedoch mindestens eine Silbe auf den Akzent folgt, wird ein tiefer Endpunkt an der unteren Grenze der Sprechstimme erreicht. Bei durchgehender Sonorität der Akzentsilbe befindet sich das Gipfelmaximum am Silbenoffset, während in einer Silbe mit Obstruenten der Maximalpunkt des Gipfels am Offset des Akzentvokals realisiert wird. Der Anstieg setzt stets zu Beginn der Akzentsilbe ein, wobei er in einer sonoren Silbe als steile F0-Bewegung realisiert wird oder bei Obstruentenumgebung als F0-Sprung zum Akzentvokal hin und steiler Anstieg auf dem Vokal verläuft.

In den als Aussagen produzierten Äußerungen sind kaum Interaktionen mit der Silben- bzw. Segmentstruktur der Äußerung zu beobachten. Das Gipfelmaximum liegt stets in der Silbe direkt vor dem Akzent. Die fallende Bewegung erstreckt sich über das sonore Silbenonset und den Akzentvokal oder wird bei einer tonalen Unterbrechung durch Obstruenten als ein minimaler F0-Sprung zu einer tieferen Frequenz, auf der die Bewegung über den Akzentvokal einsetzt, realisiert. Der Endpunkt der fallenden Gipfelbewegung wird unabhängig von der Beschaffenheit der Akzentsilbe entweder am Offset oder im letzten Drittel des Akzentvokals erreicht, im Anschluß daran ist ein leicht fallendes Auslaufen der

Satzkontur festzustellen. Falls ein Anstieg zum Gipfel realisiert wird, setzt er auf den ersten tonalen Segmenten der Äußerung ein.

### **3.2.3. Diskussion**

In der durchgeführten Voruntersuchung erweist sich in erster Linie die Positionierung des Gipfelmaximums in Relation zum Akzentvokal bzw. zur Akzentsilbe und die damit im Zusammenhang stehende Richtung der F0-Bewegung über den tontragenden Segmenten der Akzentstelle als ein sehr wichtiger, keinen Variationen unterliegender Parameter. Dieses Ergebnis, daß die zeitliche Gipfelkoordination als ein relevanter *cue* für die Kodierung der Satzmodusdistinktion im Russischen anzusehen ist, bestätigt die unter 3.1. referierte, ebenso empirisch belegte Feststellung von Makarova (1999). Auch für andere Sprachen, beispielsweise für das Deutsche, wurde die Relevanz dieses Parameters in Produktion und Perzeption empirisch fundiert (vgl. Kohler, 1991a; Gartenberg und Panzlaff-Reuter, 1991). Für das Deutsche konnte beim Gipfeltiming eine dreifache Kontrastbildung (*früh* vs. *mittel* vs. *spät*) festgestellt werden, durch welche ein Sprecher unterschiedliche Beziehungen zum Gesagten ausdrücken kann (vgl. Kohler, 1991a: 160): Durch einen frühen Gipfel kann ein Sprecher seine Kenntnis von altbekannten Fakten und Argumentationen signalisieren, während ein mittlerer Gipfel Erkenntnis und Offenheit den neuen Fakten gegenüber anzeigt und ein später Gipfel Verwunderung über unerwartete Fakten zum Ausdruck bringt. Eine Reihe von Hörexperimenten mit Sprechern vieler anderer Sprachen sowohl ohne als auch mit Deutschkenntnissen, die ein ähnliches Ergebnis zum Differenzierungsvermögen zwischen intonatorischen Mustern mit einem frühen vs. nicht-frühen Timing lieferte, veranlaßte Kohler (1991a: 156) dazu, Timing als „*a general psychophonetic, language-independent phenomenon, which may be incorporated into a language-specific phonology at different levels*“ zu interpretieren. Der intonatorische Parameter der Gipfelpositionierung wird im Russischen vs. im Deutschen unterschiedlichen funktionalen Domänen zugeordnet. In Termini von Bühler (1934) erfüllt der Timing-Unterschied im Deutschen eine Ausdrucksfunktion, während er im Russischen für die Appellfunktion eingesetzt wird. Die von Kodzasov (1996, 1999) geforderte Differenzierung des Timing-Parameters in Abhängigkeit von der zeitlichen Synchronisierung der Tonbewegung mit den Segmenten der Akzentsilbe kann für die vorliegenden Daten nicht als Notwendigkeit erkannt werden.

Ein weiterer ausgeprägter Unterschied in Fragen vs. Aussagen betrifft ihre Gipfelgestalten und scheint in erster Linie durch Gipfelhöhe und Aufstiegssteigung bestimmt zu

sein. Im Vergleich dazu erscheint der Abstieg des Gipfels für die Kontrastkodierung weniger relevant, da er durch den Unterschied in der Gipfelhöhe und der Gipfelpositionierung determiniert und in Fragen von der Silbenanzahl am Äußerungende abhängig ist.

Die Ergebnisse der Voruntersuchung zeigen auch einige Abweichungen von den aus der Literatur bekannten Beschreibungen der zu diskutierenden Intonationskonturen. Der von Bryzgunova (1980: 111) verzeichnete F<sub>0</sub>-Abstieg im Akzentvokal von Frageäußerungen wurde in den bearbeiteten Daten nicht festgestellt und kann eventuell durch die bei Bryzgunova nicht berücksichtigten mikroprosodischen Einflüsse der Konsonanten auf die Ausprägung der globalen Gipfelkontur erklärt werden. Ein F<sub>0</sub>-Plateau am Offset des Akzentvokals, welches nach Angaben von Svetozarova (1982, 1998) und auch Bryzgunova (1980) in Fragen vorkommt, wurde in dem analysierten Korpus nicht beobachtet. Die meisten Gipfelkonturen in Fragen konnten durch einen Einsatz-, Maximal- und Endpunkt der F<sub>0</sub>-Bewegung adäquat beschrieben werden, die Ausnahme bildeten die hoch endenden Konturen ohne realisierten Abstieg. Neu ist die in der vorliegenden Untersuchung registrierte Tendenz, daß die Positionierung des Gipfelmaximums in Frageäußerungen von der Sonorität der Akzentsilbe abhängt und daher nicht ausschließlich in Relation zum Akzentvokal betrachtet werden soll. Die hier festgestellte freie Variation im Vorlauf einer Aussageäußerung widerspricht der Meinung von Nikolajeva (1977: 81), der Unterschied zwischen einem ebenen Verlauf und einem Anstieg vor der Akzentsilbe eines Deklarativsatzes sei relevant. Diese Behauptung muß zumindest für die in dieser Untersuchung analysierte Art von Daten eingeschränkt werden. Andererseits liegen den Ergebnissen der hier diskutierten Voruntersuchung die Aufnahmen einer einzigen Sprecherin zugrunde, so daß weitere Datenerhebungen der gleichen Art von anderen Sprechern des (Standard-)Russischen notwendig sind, um sichere Aussagen zu den sprachlich relevanten Produktionsmustern treffen zu können.

In der nun folgenden Tabelle (Tab. 3./2.) sind die akustischen Parameter zusammengefaßt, durch welche der intonatorische Unterschied zwischen Gipfelkonturen in Fragen vs. Aussagen in der Produktion umfassend beschrieben werden kann.

Tab. 3./2.: Zusammenfassung der charakteristischen Parameterunterschiede zwischen den Grundfrequenzkonturen in Fragen vs. Aussagen anhand der durchgeführten Produktionsuntersuchung.

<i>F0-Parameter</i>	<i>Aussage</i>	<i>Frage</i>
Intervall des Gipfelanstiegs	gering (2 HT)	umfangreich (6 HT)
Neigung des Gipfelanstiegs	langsam (7 HT/s)	schnell (28 HT/s)
Gipfelhöhe	6 HT tiefer als in Fragen	6 HT höher als in Aussagen
Gipfeltiming	früh	spät

## **4. Ansätze der Beschreibung von Intonationskontrasten in der Perzeption**

In diesem Kapitel werden Untersuchungen besprochen, die intonatorische Kontraste unter perzeptivem Aspekt behandeln. Zunächst soll der aktuelle Stand der Forschung zum Russischen beleuchtet werden (Unterkapitel 4.1.). Im Abschnitt 4.2. wird auf ausgewählte Untersuchungen eingegangen, die sich mit der Fragestellung der intonatorischen Satzmodus-kodierung in anderen Sprachen auseinandersetzen. Im Unterkapitel 4.3. werden aufgrund der Ergebnisse der diskutierten Perzeptionsforschung und der unter 3.2. dargestellten Voruntersuchung in der Produktion Evidenzen für das Perzeptionssystem des Russischen zusammengefaßt.

### **4.1. Perzeptionsuntersuchungen zum Intonationssystem des Russischen**

Die wenigen z.Zt. bekannten Perzeptionsuntersuchungen zum Russischen stammen von Odé (1987, 1989), die den IPO-Ansatz der perzeptiven Intonationsanalyse ('t Hart, Collier und Cohen, 1990) auf das Russische anwendete. Den Ausgangspunkt des Ansatzes bildet die Erkenntnis, daß nicht alle Bewegungen im F0-Verlauf für die Wahrnehmung durch den Hörer wichtig sind. Ziel der Untersuchungen von Odé (1987, 1989) war es, grundlegende Typen von perzeptiv relevanten, Prominenzwahrnehmung hervorrufenden Grundfrequenzbewegungen (*types of pitch accent*) im Russischen zu erfassen. Dies geschah unter Anwendung der *close copy*-Stilisierungsmethode. Bei dieser Methode wird die Ausgangskontur auf wesentliche Grundfrequenzbewegungen reduziert, basierend auf der perzeptiven Äquivalenz zwischen dem produzierten Original und seiner stilisierten Version (Odé, 1989: 13ff.; s. auch Unterkapitel 5.2.2.1.). Daraufhin wurden die stilisierten Muster anhand des Kriteriums der perzeptiven Adäquatheit zu diskreten Typen von tonalen Akzenten gruppiert (Odé, 1989: 31ff.). Beide Arbeitsschritte wurden in Perzeptionsexperimenten verifiziert. Im Folgenden wird die daraus resultierende Klassifikation von Odé (1989) zusammengefaßt.

Auf Grundlage der Perzeptionsanalyse gelesener und spontaner Produktionsdaten von neun Sprechern des Russischen und anschließend durchgeführten Perzeptionsexperimenten wurden sieben Typen von fallenden und sechs Typen von steigenden tonalen Akzenten extrahiert. Diese unterscheiden sich in vier perzeptiv bedeutsamen Merkmalen: Intervall (*excursion*), Timing (*timing*) und Neigung (*slope*) sowie postakzentuale Tonlage (*posttonic*) der perzeptiv relevanten F0-Bewegung. Für die gesamte Beschreibung ist der sprecher-

spezifische Tonumfang zwischen der höchsten und der tiefsten produzierbaren Grundfrequenz maßgebend, der in zwei Register (*high* und *low*) unterteilt wird. Die Größen aller relevanten Parameter werden in Halbtonschritten (HT, engl. *semitone* (ST); s. Unterkapitel 2.1.2.) ermittelt.

Der Parameter des Intervalls bezeichnet in den Termini von Odé (1989: 89ff.) die Größe einer relevanten tonalen Bewegung, wobei entweder ein tatsächliches (*actual*) oder ein relatives (*relative*) Intervall ermittelt wird. Das relative Intervall, welches das Verhältnis zwischen der Endfrequenz einer Tonbewegung und der tiefsten Frequenz des gegebenen Sprechers darstellt, bildet die Lage der Endfrequenz einer Bewegung hoch oder tief innerhalb des Sprecherregisters ab. Das relative Intervall wird für steigende Akzenttypen berechnet. Infolgedessen ergibt sich eine Distinktion zwischen einem Anstieg mit einem normalen Intervall (*r, rise*) für die im tiefen Register liegenden Tonbewegungen und einem Anstieg mit einem großen Intervall (*R, Rise*) für Bewegungen mit Endfrequenzen im hohen Register. Das Intervall eines fallenden Akzenttyps (*F, Fall*) kann dagegen stärker variieren. Hierfür wird das tatsächliche Intervall berechnet, das sich auf die Größe des Abstands zwischen der Anfangs- und Endfrequenz der zu beschreibenden tonalen Bewegung bezieht. Die Lage der Endfrequenz einer fallenden Bewegung im Sprecherregister wird durch die postakzentuale Tonlage mit vier Möglichkeiten beschrieben: (1) Die perzeptiv relevante F<sub>0</sub>-Bewegung fällt und endet an der unteren Grenze der Sprechstimme (*l, low level*), (2) sie fällt, endet jedoch auf einem Niveau oberhalb der unteren Begrenzung der Sprechstimme (*nl, non-low level*) oder aber (3) weist eine hohe postakzentuale Tonlage (*h, high*) auf. Eine spezielle Art des fallenden Akzenttyps ist (4) die höher als *non-low-level* endende, äußerungsmedial fallende Tonbewegung (*f, fall*). Die steigenden Akzenttypen werden ebenfalls nach deren postakzentualen Tonlagen klassifiziert, wobei ein tiefes (*l, low*), ein mittleres (*m, middle*) und ein hohes (*h, high*) tonales Niveau sowie Abwesenheit des postakzentualen Tonabschnitts ( $\emptyset$ ) unterschieden werden.

Der Parameter des Timing wird bei Odé durch die Position des Endpunkts einer tonalen Bewegung innerhalb der Akzentsilbe definiert: Bei einem frühen Timing (-) wird der Endpunkt der Bewegung um den Vokalonset erreicht, bei einem späten Timing (+) liegt der Bewegungsendpunkt später in der Akzentsilbe. Unter Neigung versteht Odé (1987, 1989) die Geschwindigkeit der perzeptiv relevanten Grundfrequenzbewegung, welche in Halbtonschritten pro Sekunde (HT/s) erfaßt wird. Dieses Maß hängt von der Dauer und dem Intervall der Bewegung ab. Es wird zwischen einer steilen (mehr als 70 HT/s) und einer langsamen Neigung (weniger als 70 HT/s) unterschieden. Der Neigungsgrad wird jedoch nicht als ein

unabhängiges distinktives Merkmal im Russischen betrachtet (Odé, 1989: 95f.). In diesem Sinne weisen die steigenden Akzenttypen mit dem großen Intervall und dem konstant frühen Timing eine steile Neigung auf, während eine graduelle Neigung in allen anderen Akzenttypen vorkommt.

Die Abbildungen 4./1. und 4./2. enthalten schematische Darstellungen der *types of pitch accents* auf der logarithmischen Halbton-Skala. Auf die Skalierung der Zeitachse wurde verzichtet, angegeben wurde allein die Zeitdauer der Plateaus in den betreffenden Akzenttypen.

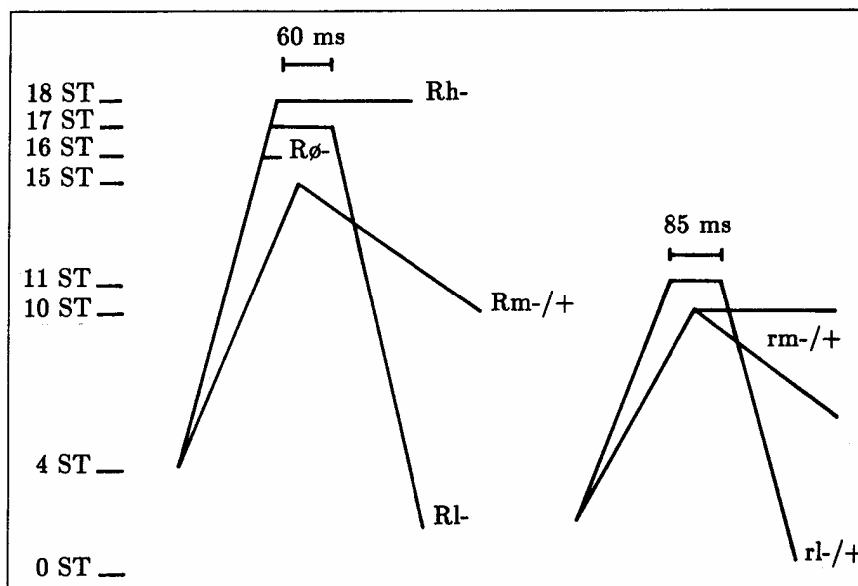


Abb. 4./1.: Typen von steigenden tonalen Akzenten (aus Odé, 1989: 120).

Für die steigenden Akzenttypen hat also das Merkmal des Timings keine differenzierende Bedeutung: Es ist entweder konstant oder kommt in freier Variation vor. Das Merkmal der Exkursion wird im Fall eines fehlenden postakzentualen Abschnitts neutralisiert. Für die steigenden Akzenttypen mit normaler Exkursion (*r*) wird die Distinktion zwischen einem Anstieg mit anschließender hoher vs. mittlerer postakzentualer Tonlage aufgehoben.

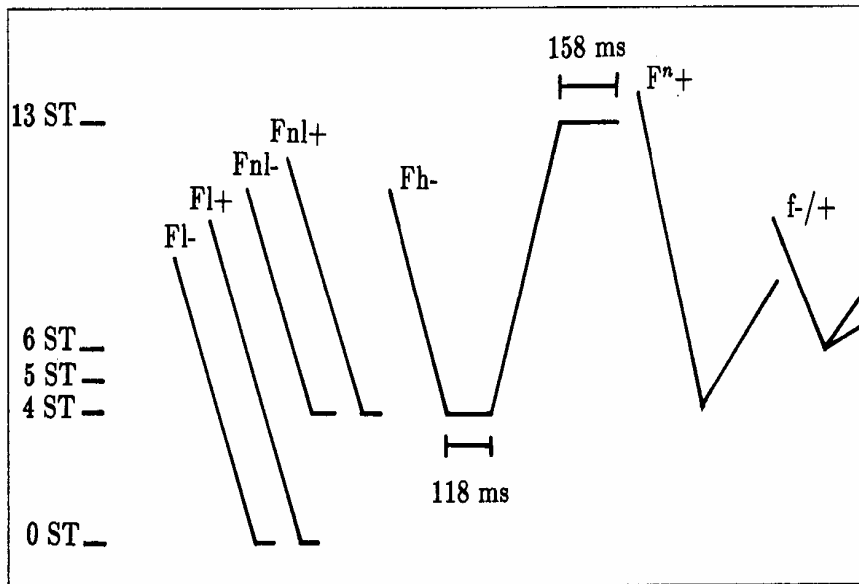


Abb. 4./2.: Typen von fallenden tonalen Akzenten (aus Odé, 1989: 125).

Bei den fallenden Akzenttypen spielt also das Timing ausschließlich für Akzente ohne anschließenden Anstieg eine differenzierende Rolle. Bei dem Akzenttyp  $F^n+$  handelt es sich um eine mehrfache Wiederholung der fallenden Tonbewegung, wobei der postakzentuale Anstieg für die Realisierung der wiederholten Konfiguration notwendig ist.

Da die linguistischen Funktionen der aufgestellten Akzenttypen bei Odé (1989: 39) unberücksichtigt bleiben, erweist sich der Vergleich mit den funktional ausgerichteten, klassifizierenden Beschreibungen auf der Basis von Produktionsdaten (s. Unterkapitel 3.1. und 3.2.) als schwierig. Kodzasov (1992: 152) stellte die formal orientierte Klassifikation von Bryzgunova (1980) der oben erläuterten Klassifikation von Odé (1989) gegenüber. Dabei wurden für die unter 3.1. diskutierten Intonationskonstruktionen von Bryzgunova folgende Parallelen aufgezeigt:

Bryzgunova (1980)	IK-1, IK-2	IK-3
Odé (1989)	Fl-, Fl+, Fnl-, Fnl+, f(?)	Rl, Rm, Rø, rl, rm

Die höhere Anzahl von steigenden und fallenden Akzenten bei Odé (1989) im Gegensatz zu Bryzgunova (1980) resultiert daraus, daß Odé (1989) akustische Parameter wie das Timing, das Intervall und die postakzentuierte Tonlage präzisierte und in das System der relevanten Merkmale integrierte.

Obwohl das Inventar an grundlegenden Intonationsformen des Russischen durch die Arbeit von Odé (1989) erweitert wurde, scheinen die relevanten Tonmuster nicht erschöpfend beschrieben zu sein: Ein für Frageäußerungen charakteristisches, steigendes Akzentmuster

mit einer zugespitzten Gipfelgestalt, welches die unter 3.2. geschilderte Produktionsuntersuchung ergab, fehlt hier. An dieser Stelle wird die Klassifikation von Odé (1989) auch von Kodzasov (1992: 153) kritisiert. Er weist darauf hin, daß ein Akzenttyp R1 ohne Plateaubildung „für manche Frageäußerungen typisch“ ist.

Die wichtige Leistung der Arbeiten von Odé (1987, 1989) ist in erster Linie die methodische Transparenz, die in den meisten, auf den Produktionsdaten basierenden Arbeiten zur russischen Intonation (s. Unterkapitel 3.1.) zu vermissen ist, und die präzise Beschreibung derjenigen akustischen F0-Parameter, auf die sich die perzeptive Distinktion der Tonbewegungen im Russischen gründet. Da jedoch die Interpretation der aufgestellten Muster durch den Hörer bei Odé (1987, 1989) nicht behandelt wird, bleibt ein wichtiger Aspekt der Sprachperzeption ungeklärt.

#### **4.2. Perzeptionsuntersuchungen zur intonatorischen Satzmodusunterscheidung in anderen Sprachen**

In diesem Abschnitt werden zwei Perzeptionsuntersuchungen dargestellt, die sich mit dem Zusammenhang zwischen akustischen Parametern und deren unterschiedlich gewichteten Auswirkungen auf die Perzeption von Satzmodi auseinandersetzen. Behandelt werden hier Sprachen, welche ähnlich wie das Russische den unter 2.2. diskutierten Satzmodusunterschied intonatorisch markieren, d.h. sowohl in Fragen als auch in Aussagen eine steigend-fallende Gipfelkontur aufweisen. Dazu gehören Ungarisch und die neapolitanische Varietät des Italienischen.

##### **4.2.1. Gósy und Terken (1994) zum Ungarischen**

Die unter 2.2.2. bereits zitierte Perzeptionsuntersuchung zum Ungarischen stammt von Gósy und Terken (1994). Die Autoren beschäftigen sich mit der Fragestellung, welche prosodischen *cues* von ungarischen Hörern bei kontextfreier Präsentation von Äußerungen benutzt werden, um zwischen Fragen und Aussagen zu differenzieren. Sie untersuchen den speziellen Fall, in dem der Satzakzent auf ein äußerungsfinales bisyllabisches Wortes fällt, so daß der relevante tonale Wechsel in beiden Satzmodi auf der Pänultima realisiert wird. In anderen Fällen, wie z.B. in Äußerungen, deren Satzakzent final auf einem Wort mit mehr als zwei Silben liegt, werden beide Satzmodi eindeutig durch abweichende Koordination der steigend-

fallenden Gipfelbewegung mit den Segmenten der Äußerung differenziert (s. auch Unterkapitel 2.2.2.).

In einem Perzeptionsexperiment analysieren Gósy und Terken (1994) drei Gipfelcharakteristika (Timing, Höhe und Gestalt des Gipfels in der Akzentsilbe) sowie Vorhandensein eines weiteren Akzents am Äußerungsanfang hinsichtlich deren Auswirkungen auf die Interpretation einer Äußerung als Aussage bzw. Frage. Ausgehend von dem als Aussage produzierten Satz *De elkérte a labdát.* (*Aber er fragte nach dem Ball.*) wurden systematische Manipulationen am Gipfeltiming vorgenommen, indem die originale steigend-fallende Gipfelkontur in Relation zum Akzentvokal in mehreren äquidistanten Schritten auf der Zeitachse verschoben wurde. Bei einer weiteren Manipulation wurde aus der Reihe jeweils ein Gipfelmaximum mit Positionierung vor dem Akzentvokal (früher Gipfel), mitten im Akzentvokal (mittlerer Gipfel) und nach der Vokaloffsetgrenze (später Gipfel) in mehreren Schritten erhöht, ohne daß die begleitende, zwangsläufige Änderung in der Gipfelgestalt berücksichtigt wurde. Um den Effekt der Gipfelgestalt zu testen, wurden der An- sowie der Abstieg in der natürlich produzierten steigend-fallenden Gipfelkontur verändert. Im Gegensatz zum steil verlaufenden Steigen und Fallen des Originals wurde ein gradueller Anstieg für den mittleren und späten Gipfel sowie ein gradueller Abstieg für den frühen und mittleren Gipfel resynthetisiert. Die letzte Manipulation betraf eine zusätzliche steigend-fallende Akzentkontur am Anfang der Äußerung (in *elkérte*). Die Höhe des zusätzlichen Gipfels variierte in Relation zum finalen Gipfel und war höher, gleich hoch oder tiefer als der letztere.

Die Ergebnisse des auf die Identifikation der Items ausgelegten Perzeptionsexperiments zeigen, daß die vier berücksichtigten Eigenschaften der Äußerungskontur die Hörerinterpretationen in unterschiedlichem Maße beeinflussen. Die Parameter des Timing und der Gipfelhöhe scheinen die stärkeren *cues* in der Perzeption des untersuchten Satzmoduskontrastes zu sein. Die Relevanz der Gipfelgestalt und eines zusätzlichen, vorangehenden Gipfels in der Äußerung ist niedriger einzustufen.

Hierbei ist festzustellen, daß mittlere Gipfel, deren Hochpunkt in der zweiten Hälfte des Akzentvokals liegt, mit einer Frage assoziiert werden. Für späte Gipfel, die ein F0-Maximum am Vokaloffset oder später in der Silbe aufweisen, läßt diese Tendenz geringfügig nach. Frühe Gipfel, deren F0-Maximum sich um den Vokaloffset befindet, werden als Aussagen beurteilt. Höhere Gipfel in jeder Position beeinflussen die Wahrnehmung einer Äußerung als Frage generell positiv, wobei dieser Effekt am stärksten für den mittleren, weniger stark für den späten und am schwächsten für den frühen Gipfel ausgeprägt ist. Es

scheint eine wechselseitige Abhängigkeit zwischen diesen beiden *cues* im Hinblick auf die perzeptive Kodierung des Satzmodusunterschieds zu bestehen, das Gipfeltiming wird jedoch stärker gewichtet als die Gipfelhöhe. Die Änderung in der Gipfelgestalt scheint eine schwächere Auswirkung auf die Perzeption der Kategorien zu haben: Mittlere und späte Gipfel mit einer graduellen Neigung des Anstiegs rufen weniger häufig die Wahrnehmung einer Frage hervor als Gipfel mit einer steilen Neigung. Bei einem graduellen Abstieg in frühen und mittleren Gipfeln dagegen wachsen die Frage-Urteile im Vergleich zur Kontur mit einem steilen An- und Abstieg. Die Präsenz eines zusätzlichen Gipfels in der Äußerung hat einen negativen Effekt auf die Wahrnehmung als Frage für alle Gipfelpositionen.

Im Ungarischen wird also der perzeptive Unterschied zwischen Frage- vs. Aussageäußerungen durch mehrere Parameter kodiert, welche vor allem die Ausprägung einer Gipfelkontur in Relation zum Akzentvokal betreffen. Die distinktiven *cues* sind hierbei frühe vs. mittlere/ späte Gipfelpositionierung, tiefe vs. hohe Gipfelhöhe, gradueller vs. steiler An- und Abstieg des Gipfels.

#### **4.2.2. D'Imperio und House (1997) über die neapolitanische Varietät des Italienischen**

Die Untersuchung von D'Imperio und House (1997) befaßt sich mit der perzeptiven Distinktion zwischen Frage und Aussage in der neapolitanischen Varietät des Italienischen. Die Autoren versuchen zu klären, welcher der zwei Parameter in der perzeptiven Kategorisierung als der stärkere *cue* fungiert: allein die zeitliche Synchronisierung des Gipfelhochpunkts mit der Akzentsilbe oder aber die steigende bzw. fallende Tonbewegung über der Akzentsilbe.

Es werden zwei Arten von tonalen Akzenten in Aussagen sowie ein Tonakzent in Frageäußerungen beschrieben, welche sich durch die Positionierung des Gipfelmaximums innerhalb der Akzentsilbe unterscheiden (s. D'Imperio und House, 1997: 251). In dieser Varietät des Italienischen sind Aussagen mit einem breiten Fokus durch ein Gipfelmaximum vor der Akzentsilbe und eine fallende tonale Bewegung über der Silbe gekennzeichnet, während Aussagen mit einem engen Fokus das Gipfelmaximum in der Akzentsilbe aufweisen. In letzteren findet der Gipfelanstieg in der Akzentsilbe statt, der Abstieg verläuft über den Akzentvokal. Der Tonakzent in Fragen wird durch einen lokalen Tonanstieg auf der Akzentsilbe charakterisiert, dessen Maximum später in der Akzentsilbe erreicht wird als in Aussagen mit einem engen Fokus.

Ausgehend von der als Aussage mit engem Fokus produzierten Äußerung „*Mamma andava a ballare da Lalla*“ (Mutter pflegte, bei Lalla tanzen zu gehen) wurden zwei Manipulationen vorgenommen: Zum einen wurde durch eine sukzessive Gipfelverschiebung ein Kontinuum mit fünf Gipfelpositionierungen über dem Akzentvokal erzeugt, zum anderen ein Kontinuum mit fünf Verschiebungen der Abstiegsbewegung über dem Akzentvokal bis in die nachfolgende unakzentuierte Silbe, wobei mit jedem Manipulationsschritt ein längeres Grundfrequenz-Plateau über dem Akzentvokal entstand. Ausgehend von der gleich lautenden, natürlich produzierten Frageäußerung wurden zwei vergleichbare Kontinua hergestellt: Das erste Kontinuum entstand durch fünf Verschiebungen des Gipfelmaximums vom Vokaloffset bis zum Vokaloffset hin, beim zweiten wurde die Anstiegsbewegung durch den Akzentvokal verschoben, bis sie über die Segmente vor dem Akzentvokal verlief und sich ein F<sub>0</sub>-Plateau über dem Akzentvokal erstreckte.

Die Ergebnisse des durchgeführten *forced-choice*-Identifikationstests mit den beschriebenen Stimuli zeigen, daß in beiden Kontinua der Gipfelverschiebung eine scharfe Kategoriengrenze vorliegt: Die Stimuli mit dem Gipfelmaximum im letzten Drittel des Akzentvokals und an der Offsetgrenze rufen das Frage-Perzept hervor, eine Aussage wird im Fall der Gipfelmaxima um das Vokaloffset wahrgenommen. Die Stimuli mit Abstiegsverschiebung ausgehend von der Aussagekontur wurden überwiegend als Aussagen beurteilt. Deren Interpretationen wurden uneindeutig, wenn der Abstieg im letzten Drittel des Akzentvokals ansetzte und über die nachfolgende Silbe verlief. Die letzte Serie, in welcher der Anstieg auf der Basis des Frageoriginals zum Vokaloffset hin verschoben wurde, löste zu mehr als 80%-Fällen eine Fragewahrnehmung aus. Die 100%-Grenze wurde erreicht, sobald der Aufstieg über die erste Hälfte des Akzentvokals erfolgte.

Die in dieser Untersuchung gewonnenen Ergebnisse wurden so interpretiert, daß in erster Linie eine steigende vs. fallende tonale Bewegung über einem akzentuierten Vokal (als einem Bereich mit spektraler Stabilität) für die perzeptive Distinktion zwischen Frage und Aussage in dieser Varietät des Italienischen verantwortlich ist, wobei eine fallende Bewegung Aussage und eine steigende Bewegung Frage indiziert. Der späte Abstieg in der postakzentuierten Silbe für Fragen und der frühe Anstieg vor dem Akzentvokal für Aussagen wurden als sekundäre perzeptive *cues* gesehen, wie auch die Gipfelhöhe und die Anstiegsneigung.

Eine solche Interpretation scheint jedoch die Komplexität der Parameterinteraktion zu vereinfachen. Die Ergebnisse für die Randstimuli der getesteten Kontinua, in denen der Abstieg der Kontur von Aussageäußerung vs. der Aufstieg der Kontur von Frageäußerung außerhalb des Akzentvokals liegen und sich ein hohes Plateau über dem Akzentvokal

erstreckt, blieben bei der Diskussion vernachlässigt. Diese Ergebnisse zeigen aber deutlich, daß die beiden ambig erscheinenden Stimuli, die sich ausschließlich durch die Anstiegsneigung unterscheiden, trotzdem in über 80% bis 100% der Fälle als eine Frage bzw. eine Aussage identifiziert werden konnten. Diese Ergebnisse legen nahe, daß nicht nur der F<sub>0</sub>-Verlauf über dem Akzentvokal sondern auch die Anstiegsneigung als ein bedeutsamer *cue* zu interpretieren ist. Diese Neigung, die in Gipfelkonturen von Aussagen steiler ist als in Konturen von Fragen, scheint auch über dem Bereich mit spektralen Veränderungen wie in *da Lalla* perzipierbar zu bleiben. Spektrale Diskontinuität im Sinne von House (1990) bezieht sich auf stimmlose oder nasale Abschnitte im Signal, die auch von einem Einbruch in der Intensität begleitet werden. Dies trifft allerdings im Fall des Sonoranten /l/ in „*Lalla*“ nicht zu. In den für die Stimulusherstellung verwendeten Äußerungen weist der Lateral in der Akzentsilbe eine mit Vokalen vergleichbare Intensitätsstruktur auf.

### **4.3. Evidenzen für das Russische**

Die unter 4.2.1. und 4.2.2. erläuterten Untersuchungen zeigen, daß semantisch-pragmatische Sprachkonzepte durch ein hierarchisches Zusammenspiel von mehreren kontrastbildenden akustischen Parametern der Grundfrequenz kodiert werden können, wobei es einige dominante perzeptive *cues* sowie sekundäre, die Unterscheidung unterstützende *cues* zu differenzieren gilt. Für das Russische liegen z. Zt. keine vergleichbaren Untersuchungen vor. Die referierte Arbeit von Odé (1989) gibt Hinweise darauf, daß sich die perzeptive Unterscheidung zwischen Fragen und Aussagen im Russischen unter anderem auf solchen intonatorischen Merkmalen wie Intervall, Timing und Endlage der prominenten Gipfelbewegung gründet. Mit Blick auf die Ergebnisse der unter 3.2. geschilderten Voruntersuchung kann angenommen werden, daß für die Kontrastbildung im Russischen auch der Parameter der Anstiegsneigung bedeutsam ist. Eine Unterscheidung in tatsächliches und relatives Intervall, wie sie sich bei Odé (1989) findet, erscheint für die Fragestellung der vorliegenden Arbeit weniger sinnvoll als die Differenzierung zwischen dem tatsächlichen Intervall der Gipfelbewegungen und der Gipfelhöhe. Die Begriffsausdrücke „*Timing*“ und „*Gipfelposition*“ werden im Folgenden synonym gebraucht.

## **5. Empirische Untersuchung zur Perzeption der Frage im Russischen**

### **5.1. Hypothesen**

Das Ziel der durchzuführenden Perzeptionsuntersuchung ist also, eine Beziehung zwischen den akustischen Parameterausprägungen und der Perzeption von semantisch-pragmatischen Sprachkonzepten aufzuzeigen. In der vorliegenden Arbeit sollen drei akustische Grundfrequenzparameter hinsichtlich ihrer perzeptiven Relevanz für die Kodierung des Unterschieds zwischen Frage und Aussage im Russischen untersucht werden: (1) zeitliche Positionierung und (2) Höhe des Gipfels sowie (3) Neigung des Gipfelaufstiegs. Folgende *vier Hypothesen* sollen getestet werden:

1. Die oben diskutierten Untersuchungen zum Russischen und zu anderen Sprachen legen nahe, daß die Positionierung des Gipfelmaximums in Relation zum Akzentnukleus als der primäre perzeptive *cue* anzunehmen ist. Dieser Hypothese zufolge soll ausschließlich durch die Variation in der Gipfelpositionierung eine Änderung des semantisch-pragmatischen Konzeptes erzielt werden, wobei die Äußerungen mit einer Gipfelposition spät in der Akzentsilbe als Fragen perzipiert werden.

2. Die Ergebnisse der unter 3.2. geschilderten Voruntersuchung zur Produktion von Fragen und Aussagen im Russischen weisen darauf hin, daß die Höhe des Gipfels als der sekundär wichtige *cue* für die Wahrnehmung der Frage vs. Aussage einzuschätzen ist. Demzufolge soll ein höherer Gipfel in Verbindung mit einer späten Gipfelposition die Wahrnehmung der Frage deutlich verbessern.

3. Ein mit dem Parameter der Gipfelhöhe konkurrierender, sekundärer *cue* ist die Neigung des Gipfelanstiegs (vgl. Unterkapitel 3.2.). Falls dieser Parameter im Vergleich zur Gipfelhöhe dominiert, sollen bei der späten Position eines steil ansteigenden Gipfels häufiger *Frage*-Urteile auftreten als bei einem hohen Gipfel mit einem flach verlaufenden Anstieg.

4. Falls die drei extrahierten Parameter einer Gipfelkontur (s. oben) die perzeptive Unterscheidung zwischen Frage und Aussage im Russischen erschöpfend beschreiben, soll die Beurteilung eines hohen, spät in der Akzentsilbe liegenden Gipfels mit einem steilen Anstieg mit der des Originals übereinstimmen. Ein anderes Ergebnis würde darauf hindeuten, daß die intonatorische Kodierung des hier diskutierten Satzmodusunterschieds durch die drei

F0-Parameter nur unzureichend beschrieben ist und weitere perzeptive *cues* differenziert werden müssen.

## **5.2. Methode**

### **5.2.1. Sprachmaterial**

Für das Experiment wurde der Satz „*Jejo zovut Jeliena*“ (/ji'jo za'vut ji'l'jɛna/) ausgewählt, der sowohl als Aussage („*Sie heißt Elena.*“) als auch als Frage („*Heißt sie Elena?*“) produziert werden kann. Der Satzakzent wird in beiden Fällen äßerungsfinal auf der Pänultima realisiert. Die Segment- und Silbenstruktur der Äußerung scheinen sich für die Zwecke einer intonatorischen Resynthese, wie sie für das vorliegende Experiment geplant wurde, zu eignen. Der Akzentvokal befindet sich in einer sonoren Umgebung, in welcher die für die vorliegende Untersuchung beabsichtigten Manipulationen am F0-Gipfel wahrnehmbar bleiben sollen. Da die Äußerung sowohl vor als auch nach dem Akzent mindestens eine weitere Silbe aufweist, soll das segmentelle Material für eine Resynthese von Gipfelbildungen mit Auf- und Abstiegsbewegungen ausreichen. Der Vorteil dieser als Aussage sowie als Frage realisierbaren Äußerung besteht darin, daß sie aus dem realen alltagssprachlichen Gebrauch stammt. Ihre lexikalische Form und kommunikative Semantik klingen nicht konstruiert, sondern natürlich.

Eine Äußerung, die aus einem durchgehend stimmhaften segmentellen Material besteht, wäre für die Zwecke der F0-Manipulationen zweifelsohne ideal geeignet und im Russischen sicherlich konstruierbar, z.B. „*Mama lomala malinu.*“ (dt. „*Die Mutter brach den Himbeerstrauch*“). Eine solche Äußerung wäre jedoch in der natürlichsprachlichen Alltagskommunikation unakzeptabel, könnte auf naive Hörer irritierend wirken und das Verständnis resynthetisierter Sprache erschweren. Aus diesen Gründen lassen sich konstruierte Sätze nur bedingt für Perzeptionsuntersuchungen verwenden.

Der für die vorliegende Untersuchung gewählte Satz „*Jejo zovut Jeliena*“ wurde in zwei verschiedene Kontexte eingebettet, welche die Akzentstruktur und den Modus der zu realisierenden Äußerungen eindeutig determinieren mußten. Dabei bezog sich die Zieläußerung auf das Thema des vorangehenden Satzes in einer erweiternden, präzisierenden Weise.

Die als Aussage zu realisierende Äußerung wurde mit folgendem Kontextsatz versehen:

russ. *Eto - nasha novaja sosiedka. (Jejo zovut Jeliena.)*

dt. Das ist unsere neue Nachbarin. (Sie heißt Elena.)

Die Äußerung, die als Frage produziert werden mußte, erhielt folgenden einleitenden Satz:

russ. *Kak zovut tvoju podrugu? (Jejo zovut Jeliena?)*

dt. Wie heißt deine Freundin? (Heißt sie Elena?)

Die Sätze wurden von einer Muttersprachlerin des Standardrussischen (Autorin dieser Arbeit) unter Studiobedingungen vorgelesen und aufgenommen. Die digitalisierten Sprachsignale wurden nach einem wortphonematischen Transkriptionsprinzip (vgl. Bondarko 1998: 18) unter dem Sprachverarbeitungsprogramm *xassp* (URL1) segmentiert.

Die Grundfrequenzkonturen der produzierten Originaläußerungen weisen mehrere formale Differenzen auf, welche die Position und Höhe des Gipfels, die Neigung des Auf- und Abstiegs sowie den Endpunkt der Gesamtkontur betreffen (vgl. Abb. 5./1. und Abb. 5./2.). Der Gipfel in der intendierten Aussage liegt vor der Akzentsilbe, die F0-Bewegung über dem akzentuierten Vokal ist fallend. In der als Frage produzierten Äußerung wird der Gipfel dagegen spät im Akzentvokal erreicht, hier liegt eine steigende F0-Bewegung über dem akzentuierten Vokal vor. Das Gipfelmaximum liegt höher in der Frageäußerung. In diesem Fall erfolgt der Aufstieg sowie der Abstieg des Gipfels schnell über zwei Silben. Dagegen erstreckt sich die Gipfelbildung in der Aussage über mehrere Silben und ist weniger ausgeprägt. Die gesamte Äußerungskontur endet in der Frage geringfügig höher als bei der Aussage. Die Segmentdauern waren in beiden Äußerungen annähernd gleich.

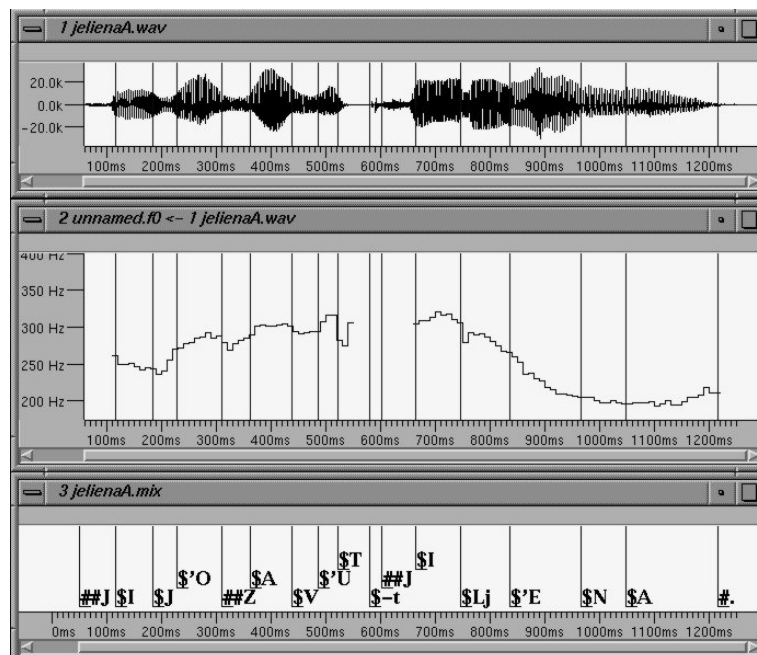


Abb. 5./1.: Sprachsignal, Grundfrequenzkontur und wortphonematisch basierte, segmentelle Etikettierung der natürlich produzierten Aussage „Jejo zovut Jeliena.“

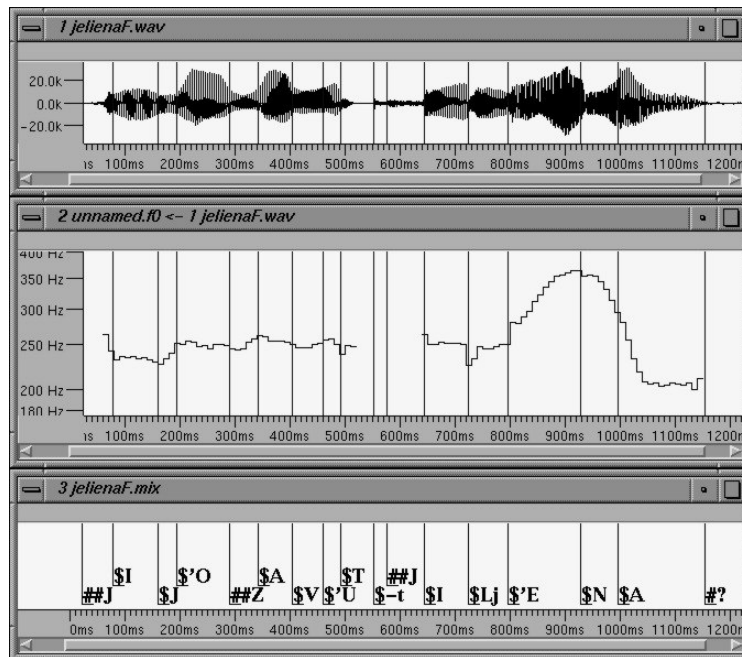


Abb. 5./2.: Sprachsignal, Grundfrequenzkontur und wortphonematisch basierte, segmentelle Etikettierung der natürlich produzierten Frage „Jejo zovut Jeliena?“

## 5.2.2. Anfertigung der Experimentstimuli

Zur Herstellung der Stimuli für das Experiment wurde das PSOLA-Verfahren (*Pitch Synchronous Overlap and Add*, s. Moulines und Charpentier, 1990) angewendet, das die Manipulationen eines Sprachsignals im Zeit-, Grundfrequenz- und Intensitätsbereich mit Beibehaltung der authentischen Stimmqualität des Sprechers ermöglicht. Da hierbei die originalen akustischen Charakteristika der jeweiligen Sprecherstimme unverändert bleiben, eignet sich PSOLA besonders gut zur Resynthese von Stimuli für Perzeptionsexperimente. Für die erforderlichen Grundfrequenz-Manipulationen und die Resynthese wurde das Signalverarbeitungsprogramm *praat* (URL2) verwendet.

### 5.2.2.1. Stilisierung der Originalkonturen

Die natürlich produzierten Konturen der beiden Äußerungen wurden nach der IPO-Methode ('t Hart et al., 1990; Adriaens, 1991; Odé, 1989) mit 5 relevanten Punkten stilisiert. Jeweils ein Punkt markierte den Anfang und das Ende des globalen Grundfrequenzverlaufs über der Äußerung. Drei Punkte beschrieben die Gipfelkontur, die durch den Einsatz der Aufstiegsbewegung, den Hochpunkt und das Ende der Abstiegsbewegung definiert wurde. Bei dieser

Art von Stilisierung werden natürliche F0-Kurven durch gerade Linien zwischen wenigen wesentlichen Punkten von mikroprosodischen Einflüssen abstrahiert und so auf sprachlich relevante Grundfrequenzänderungen unter Berücksichtigung der perzeptiven Äquivalenz mit dem Original reduziert. Die Vorgehensweise der Stilisierung stellte unter *praat* eine lineare Interpolation zwischen den fünf relevanten Punkten auf linearer Frequenzskala dar.

Die Abbildung 5./3. zeigt die stilisierte Kontur der Aussageäußerung im Vergleich zum Original. Die entsprechenden Stilisierungsdaten sind in der Tabelle 5./1. zusammengetragen. Der erste Punkt liegt bei 250 Hz am Einsatz der globalen Äußerungskontur. Der zweite Punkt bei 250 Hz legt den Anfang der Aufstiegsbewegung zum Gipfelhochpunkt fest. Der Aufstieg erfolgt über 525 ms, seine Steigung beträgt ca. 133 Hz/s (8,1 HT/s). Der Punkt des Gipfelmaximums liegt bei 320 Hz und wird in der Mitte des Vokals vor der akzentuierten Silbe als dritter Stilisierungspunkt markiert. Im letzten Drittel des 130 ms langen Akzentvokals wird der Endpunkt der Gipfelkontur bei 205 Hz erreicht, wo dann der vierte Punkt der Stilisierung gesetzt wird. Ab dem Gipfelendpunkt fällt die Kontur flach aber kontinuierlich ab und erreicht den untersten Bereich der Sprechstimme, auf den sich der fünfte Punkt bei 195 Hz bezieht. Die Äußerungsdauer beträgt 1,17 s.

Tabelle 5./1.: Akustische Parameter der Stilisierung der natürlich produzierten Aussageäußerung.

<i>Strecke</i>	<i>Dauer</i> (ms)	<i>Anfangs-F0</i> (Hz)	<i>End-F0</i> (Hz)	<i>Exkursion</i>		<i>Steigung</i>	
				<i>Hz</i>	<i>HT</i>	<i>Hz/s</i>	<i>HT/s</i>
1.-2.	130	250	250	0	0	0	0
2.-3.	525	250	320	170	4,3	133	8,1
3.-4.	245	320	205	-115	-7,7	-469	-31,5
4.-5.	270	205	195	-10	-0,9	-37	-3,2

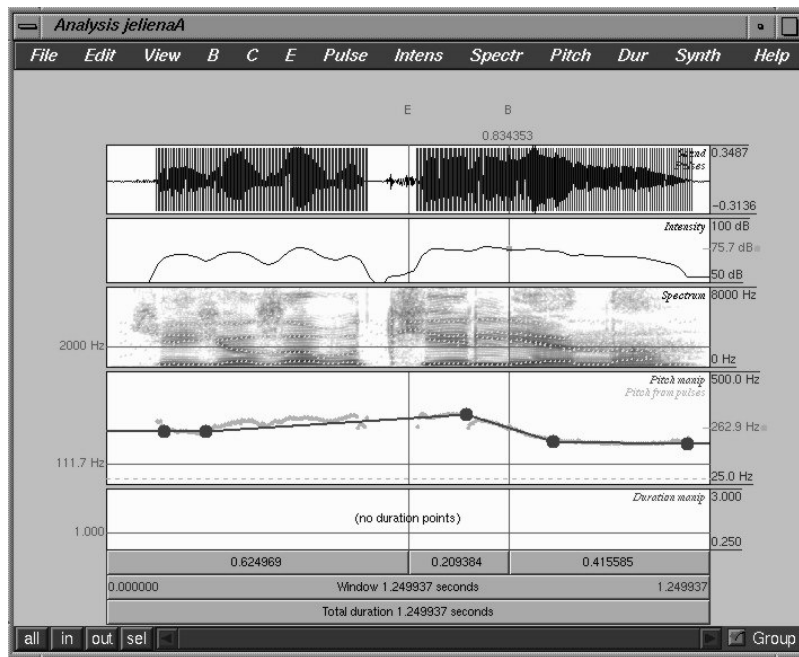


Abb. 5./3.: Oszillogramm mit Periodenmarkierungen, Intensitätskurve, Spektrogramm der natürlich produzierten Aussage „Jejo zovut *Jeliena*“. Das Grundfrequenz-Analysefenster stellt die Originalkontur (hellgraue Punkte) im Vergleich zur Stilisierungskontur (schwarze Linien) dar. Die vertikale Linie B markiert den Beginn des Akzentvokals in „*Jeliena*“.

In der stilisierten Kontur der Frage liegen die ersten beiden Punkte zur Markierung des Einsatzes der globalen Kontur und des Gipfels gleich hoch wie bei der Aussage (250 Hz). Der Gipfelanfangspunkt wird auf der Zeitachse anders plaziert und liegt an der Grenze zwischen dem Lateral und dem Vokal der Akzentsilbe. Der Aufstieg zum Gipfel hin erfolgt über dem Akzentvokal, dessen Dauer 130 ms beträgt. Das Gipfelmaximum befindet sich an der Grenze zwischen dem Akzentvokal und dem nachfolgenden Nasal und ist höher als bei der Aussage: Hier liegt der dritte Stilisierungspunkt bei 365 Hz. Die Steilheit des Aufstiegs zum Gipfel hin beträgt ca. 885 Hz/s (50,4 HT/s). Der Endpunkt der Gipfelkontur (der vierte Punkt in der Stilisierung) liegt bei 208 Hz und wird im ersten Drittel des Vokals der Endsilbe erreicht. Der Endpunkt der Kontur liegt bei 205 Hz, also etwas höher als bei der Aussage. Die Äußerungsdauer beträgt 1,135 s. Auf der Abbildung 5./4. ist die Stilisierung im Vergleich zur natürlichen F0-Kontur der Frage-Äußerung dargestellt. Die Tabelle 5./2. gibt den Überblick über die genauen Parameter der stilisierten Kontur.

Tabelle 5./2.: Akustische Parameter der Stilisierung der natürlich produzierten Frageäußerung.

Strecke	Dauer (ms)	Anfangs- F0 (Hz)	End-F0 (Hz)	Exkursion		Steigung	
				Hz	HT	Hz/s	HT/s
1.-2.	770	250	250	0	0	0	0
2.-3.	130	250	365	115	6,6	885	50,4
3.-4.	120	365	208	-157	-9,7	-1308	-81,1
4.-5.	115	208	205	-3	-0,3	-26	-2,2

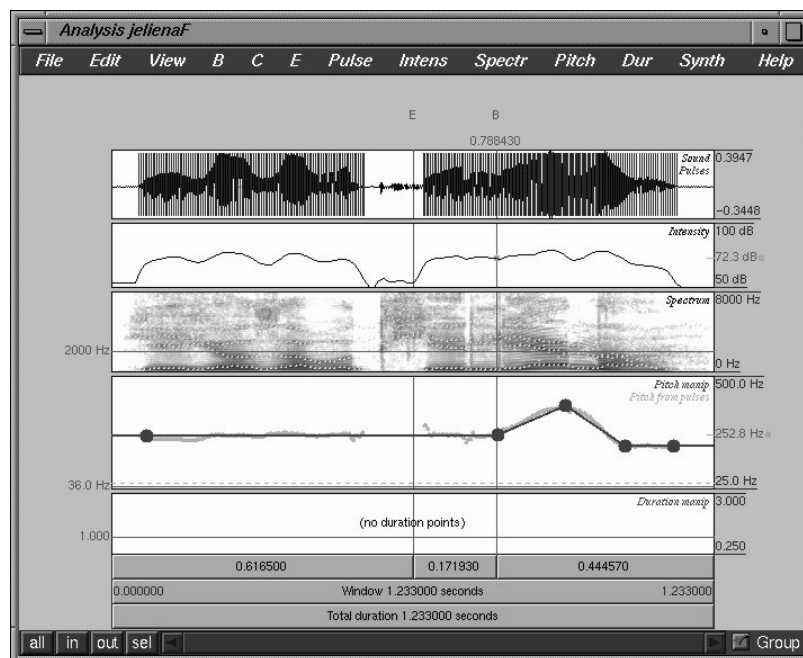


Abb. 5./4.: Oszillogramm mit Periodenmarkierungen, Intensitätskurve, Spektrogramm der natürlich produzierten Frage „Jejo zovut *Jeliena?*“. Das Grundfrequenz-Analysefenster stellt die Originalkontur (hellgraue Punkte) im Vergleich zur Stilisierungskontur (schwarze Linien) dar. Die vertikale Linie B markiert den Beginn des Akzentvokals in „*Jeliena*“.

Die erforderliche perzeptive Äquivalenz zwischen Original und Stilisierung (s. z.B. Odé, 1989: 18) ist in beiden Fällen gegeben, wurde allerdings ausschließlich durch die Autorin dieser Arbeit beurteilt und nicht in zusätzlichen Tests mit anderen Muttersprachlern verifiziert.

### 5.2.2.2. Manipulationsdesign

Die stilisierte Originalkontur der Aussageäußerung wurde der Herstellung von sämtlichen Stimuli für die vorliegende Untersuchung zugrunde gelegt. Die stilisierte Frageäußerung

wurde ohne weitere Veränderungen als ein Kontrollstimulus in das Experiment aufgenommen. Um die unter 5.1. angeführten Hypothesen zu testen, waren vier Stimulusreihen herzustellen. Diese entstanden nach einem Design, das in der Tabelle 5./3. schematisch dargestellt wird.

Die Begriffe „*Aussage*“ und „*Frage*“ in der Tabelle beziehen sich auf die Parameterausprägungen, die aus der natürlich produzierten Äußerung im jeweiligen Satzmodus extrahiert wurden, wobei die Gipfelhöhe entweder *tief* (320 Hz) oder *hoch* (360 Hz) sowie die Neigung des Anstiegs entweder *flach* (133 Hz/s) oder *steil* (885 Hz/s) sein können. Die verschiedenen Kombination der Parameter Gipfelhöhe und Aufstiegssteigung ergeben vier differente Gipfelgestalten, deren Positionierung auf der Zeitachse in einer Richtung (links - rechts) kontinuierlich manipuliert wird. Im Folgenden sollen die einzelnen Manipulationen detailliert erläutert werden.

Tabelle 5./3.: Schematischer Überblick über die Systematik der Parameterkombinationen für die Herstellung von unterschiedlichen Gipfelgestalten der vier Stimulusreihen.

<i>Stimulusreihe</i>	<i>Gipfelhöhe</i>	<i>Aufstiegssteigung</i>
1.	Aussage (tief)	Aussage (flach)
2.	Frage (hoch)	Aussage (flach)
3.	Aussage (tief)	Frage (steil)
4.	Frage (hoch)	Frage (steil)

### 5.2.2.3. Herstellung der Stimulusreihen

In der *ersten Stimulusreihe* sind die Gipfelhöhe und die Neigung des Aufstiegs ausgehend von der Kontur der natürlich produzierten Aussage unverändert gelassen, d.h. die Gipfelgestalt der Testkontur ist *tief/flach* (320 Hz/ 133 Hz/s). Die Manipulation betraf ausschließlich die Gipfelposition, die in 30-ms-Schritten über eine Strecke von 300 ms von der Mitte des Vokals vor der Akzentsilbe (die extrem frühe Gipfelposition, bei -150 ms in Relation zum Onset des Akzentvokals) bis in die Mitte des Nasals nach dem Akzentvokal (die extrem späte Position, bei +150 ms mit dem gleichen Relationspunkt) verschoben wurde (s. Abb. 5./5.).

Der Gipfelabstieg wurde ebenso unverändert verschoben, was jedoch zwangsläufig eine Umgestaltung der Endkontur bewirkte: Der 4. und der 5. Punkt der stilisierten Kontur (s.

oben) näherten sich einander bei jedem Manipulationsschritt an, bis sie beim neunten sowie zehnten Stimulus zusammenfielen. Das Vorgehen der sukzessiven Gipfelverschiebung ergab eine Gesamtzahl von 11 Stimuli.

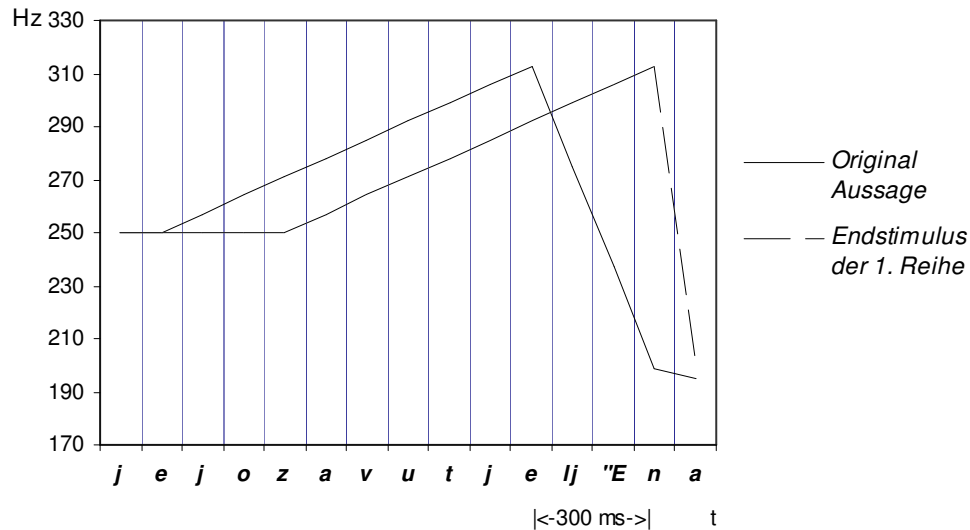


Abb. 5./5.: Schematische Darstellung der stilisierten Originalkontur der Aussage, welche die Testkontur *tief/flach* der ersten Stimulusreihe bildete, im Vergleich zum letzten Stimulus. Der Bereich, über den die sukzessiven Gipfelverschiebungen vorgenommen wurden, ist auf der Zeitachse markiert.

Die erste Stimulusreihe, die zum Testen der Hypothese zur perzeptiven Relevanz der Gipfelpositionierung eingesetzt wurde, enthielt samt dem Kontrollstimulus 12 Äußerungen.

Durch **die zweite Stimulusreihe** wurde die Hypothese zur perzeptiven Relevanz der Gipfelhöhe getestet (s. Unterkapitel 5.1.). Im Unterschied zur ersten Stimulusreihe wurde hier das Gipfelmaximum entsprechend der Höhe des Gipfels in der natürlich produzierten Frageäußerung auf 360 Hz erhöht. Damit in diesem Fall die Neigung des Gipfelaufstiegs unverändert blieb, wurde der erste Punkt des Gipfels folgerichtig 80 ms früher gesetzt. Beim Gipfelabstieg wurde zwischen dem Maximumspunkt und dem 4. Stilisierungspunkt linear interpoliert. Die auf diese Weise entstandene Testkontur *hoch/flach* ist in der Abbildung 5./6. schematisch dargestellt. Die Positionierung des Gipfelmaximums der manipulierten Kontur auf der Zeitachse wurde wie in der ersten Stimulusreihe in 30-ms-Schritten sukzessiv verändert, so daß weitere 10 Stimuli entstanden. Die gesamte Reihe umfaßte samt den zwei Referenzäußerungen 13 Stimuli.

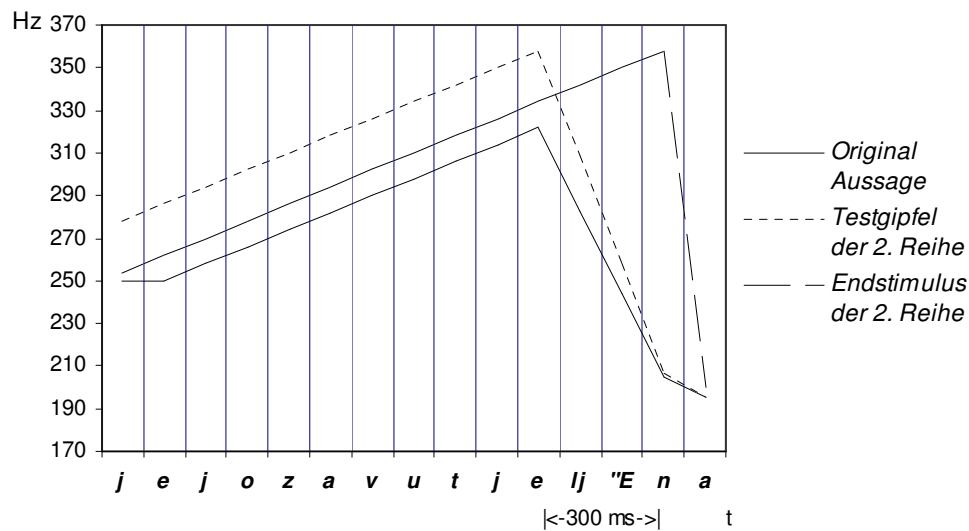


Abb. 5./6.: Schematische Darstellung der Gipfelkontur *hoch/flach*, die für die zweite Stimulusreihe angefertigt wurde, im Vergleich zum stilisierten Original der Aussageäußerung und zum letzten Stimulus. Der Bereich, über den die sukzessiven Gipfelverschiebungen vorgenommen wurden, ist auf der Zeitachse markiert.

Um den Effekt der Neigung des Gipfelaufstiegs zu testen, wurde **die dritte Stimulusreihe** hergestellt. Die Gipfelhöhe der Kontur in dieser Reihe behielt den gleichen F0-Wert wie der Gipfel in der Originalkontur der Aussage (320 Hz). Die Neigung des Gipfelaufstiegs wurde im Vergleich zu den Parametern der Aussagekontur erhöht (885 Hz/s statt 133 Hz/s, vgl. 5.2.2.2.). Der Gipfelabstieg blieb bei der Manipulation unverändert. Diese steil ansteigende Gipfelkontur (*tief/steil*) ist im Vergleich zur Ausgangskontur des Originals in der Abbildung 5./7. graphisch dargestellt. Anschließend wurde die Position des Gipfelmaximums der hergestellten Gipfelkontur in gleichen Manipulationsschritten wie in der ersten und der zweiten Stimulusreihe über der Äußerung verschoben. Die gesamte Reihe bestand aus den 11 manipulierten Stimuli und den 2 stilisierten Originaläußerungen.

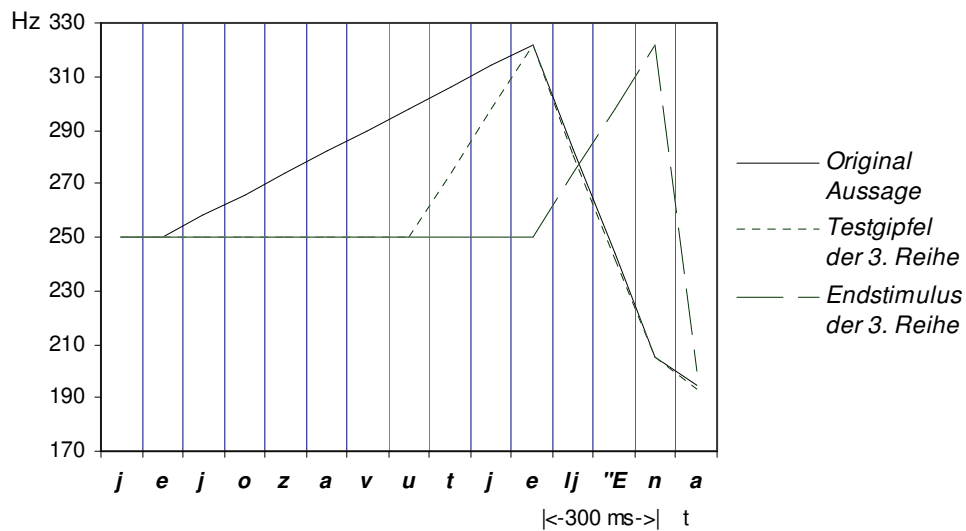


Abb. 5./7.: Schematische Darstellung der Gipfelkontur mit der Gestalt *tief/steil*, wie sie für die dritte Stimulusreihe hergestellt wurde, im Vergleich zum stilisierten Original der Aussageäußerung und zum letzten Stimulus der Reihe. Der Bereich, über den die sukzessiven Gipfelverschiebungen vorgenommen wurden, ist auf der Zeitachse markiert.

**Die vierte Stimulusreihe** zielt darauf ab, die vierte Hypothese dieser Untersuchung (s. Unterkapitel 5.1.) zu testen. Ausgehend von der stilisierten Aussagekontur wurden sowohl die Gipfelhöhe (320 Hz auf 360 Hz) als auch die Neigung des Aufstiegs zum Gipfel (von 133 Hz/s zu 885 Hz/s) geändert. Beim Gipfelabstieg wurde zwischen dem Punkt des Gipfelmaximums und dem 4. Stilisierungspunkt linear interpoliert. Die Abbildung 5./8. veranschaulicht das Schema der auf diese Weise entstandenen Kontur mit der Gestalt *hoch/steil* im Vergleich zum stilisierten Aussageoriginal. Die hergestellte Gipfelkontur wurde in 30-ms-Schritten über die Strecke von 300 ms wie in den vorangehenden Stimulusreihen verschoben. Die komplette Stimulusreihe enthielt samt den beiden stilisierten Originalen 13 Äußerungen.

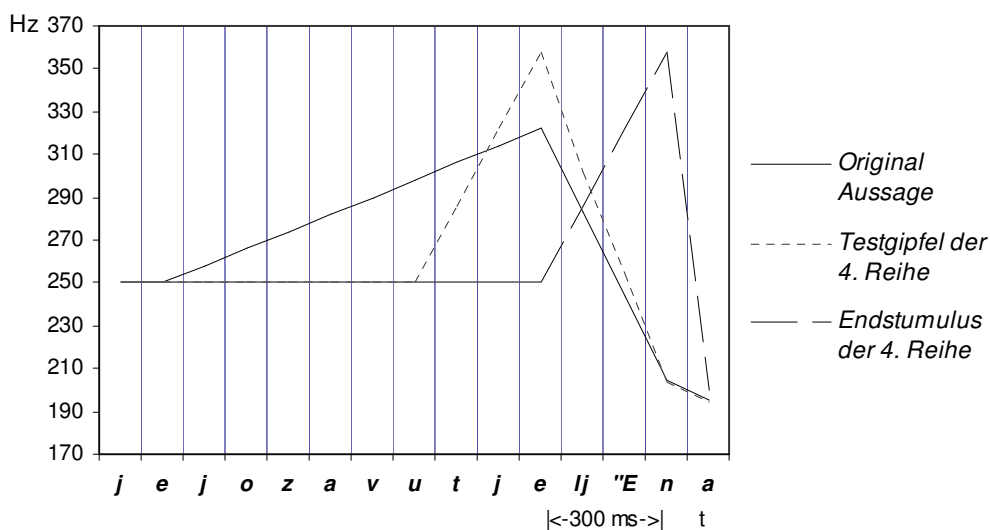


Abb. 5./8.: Schematische Darstellung der Gipfelkontur mit der Gestalt *hoch/steil*, die in der vierten Stimulusreihe resynthetisiert wurde, im Vergleich zum stilisierten Original der Aussageäußerung und zum letzten Stimulus der Reihe. Der Bereich, über den die sukzessiven Gipfelverschiebungen vorgenommen wurden, ist auf der Zeitachse markiert.

#### 5.2.2.4. Anmerkungen zu den durchgeführten Manipulationen

Bei den durchgeführten F0-Manipulationen zeigten sich zwei Nachteile des für den Versuch gewählten Satzes: Erstens erwies sich, daß lediglich eine sonore Silbe nach dem Akzentvokal nicht ausreichend segmentelles Material für die geplanten Gipfelverschiebungen lieferte. In allen Stimulusreihen fiel bei den letzten beiden Stimuli der Endpunkt der Gipfelkontur mit dem Endpunkt der Äußerungskontur auf der Zeitachse zusammen. Dies wirkte sich sowohl auf die Steilheit des Gipfelabstiegs aus, die zunahm, als auch auf die Höhe der Endkontur, die anstieg. Es handelte sich allerdings um sehr geringfügige Veränderungen (10 Hz bzw. 0,9 HT; vgl. Tab. 5./1.). Aus diesem Grund und da die betroffenen Parameter als irrelevant für den untersuchten Satzmodusunterschied angesehen wurden, konnte dieser Nachteil vernachlässigt werden.

Des weiteren enthielt die phonetische Realisierung des Testsatzes wenig Stimmhaftigkeit vor der Akzentsilbe, so daß der schnelle Gipfelanstieg im ersten Stimulus der dritten Stimulusreihe über der stimmlos produzierten Sequenz [tç] in „zovu-t J-eliena“ erfolgte, was zu einem tonalen Sprung statt einer tonalen Aufstiegsbewegung führte. Dies sollte

jedoch die Wahrnehmung des intendierten Musters nicht beeinträchtigen, da zwei unterschiedliche tonale Perzeptionsmechanismen angenommen werden können, die in Abhängigkeit von spektralen Charakteristika fungieren (House, 1990: 58ff.): In Umgebung mit spektraler Stabilität werden Tonbewegungen wahrgenommen, während sich die Perzeption in Bereichen mit spektraler Diskontinuität an Tonebenen orientiert.

### **5.2.3. Vorbereitung der Experimente**

Die oben dargestellten Stimulusreihen werden jeweils einem Perzeptionsexperiment zugrunde gelegt, das aus drei verschiedenen Tests besteht. Die im folgenden beschriebenen Tests sind in phonetischen Untersuchungen zur kategorialen Wahrnehmung etabliert, welche traditionell als eine Korrelation zwischen Diskriminations- und Identifikationsverhalten der Hörer definiert ist (vgl. 2.1.5.). Das übliche *forced-choice*-Verfahren beim Identifikationstest wird für die durchzuführenden Experimente unter anderem mit Blick auf die Kritik des Begriffes *Kategorialität* im Bereich der Intonation durch Schiefer und Baltiner (1988) abgelehnt. Stattdessen wird für die Datenerhebung eine Skala mit fünf Identifikationsitems (*eindeutige Aussage – eher Aussage – ich weiß nicht – eher Frage – eindeutige Frage*) konzipiert: Da beim *forced-choice*-Verfahren nicht von einem klaren Urteilsverhalten der Versuchspersonen ausgegangen werden kann, ist bei einer Skala mit der Möglichkeit, die Eindeutigkeit der Stimulusäußerung einzuschätzen, eine geringere Wahrscheinlichkeit eines Zufallsfehlers anzunehmen.

Ein *einfacher serieller Diskriminationstest* stellt einen Vortest dar, in dem es um eine Diskrimination der Stimuli in der Serie und eine durch Sprachintuition geleitete semantische Interpretation der diskriminierten Stimuli geht. Bei diesem Test wird eine sequentielle Aufeinanderfolge von 12 bzw. 13 Stimuli einer Reihe (wobei das stilisierte Original der Aussage jeweils den ersten und das Original der Frage den letzten Stimulus darstellen) mit einer 250 ms langen Pause zwischen den Äußerungen präsentiert. Testanfang und -ende werden durch einen 500 ms langen Piepton signalisiert. Für die erste Stimulusreihe beträgt die Testdauer 5 min 50 s, für die drei weiteren Stimulusreihen je 6 min 30 s.

In einem *Identifikationstest* soll geprüft werden, auf welche Weise die perzeptiven Unterschiede zu den in dieser Arbeit untersuchten Kategorien zugeordnet werden und wie systematisch dies geschieht. Für diesen Test werden die randomisierten Stimuli 10 mal kontextfrei mit folgender Systematik dargeboten:

ein 500 ms langer Piepton

2,5 s Pause

Stimulus 1

3,5 s Pause für die Entscheidung

Stimulus 2

...

Stimulus 11 bzw. 12

3,5 s Pause für die Entscheidung

Nach jedem 10er Block wird ein kurzer doppelter Piepton zur Orientierung gegeben. Am Anfang und am Ende des Tests ertönen 2 längere Pieptöne als Start- bzw. Abschlußsignal. Die Gesamtdauer des Identifikationstests beträgt für die erste Stimulusreihe 10 min 12 s (insgesamt 120 Stimuli), für die anderen Reihen je 11 min 2 s (130 Stimuli).

Ein *randomisierter AX-Diskriminationstest* zielt darauf ab, die perzeptiven Kategoriengrenzen in einem akustischen Kontinuum sukzessiv manipulierter Stimuli abzustechen. Zum einen werden für diesen Test jeweils 11 Stimuli pro Stimulusreihe verwendet und mit der Schrittbreite 2 (d.h. 1/3; 2/4 usw.) gepaart. Zum anderen werden auch physikalisch identische Kontrollstimuli (Schrittbreite 0) aufgenommen, wofür alle ungeradzahligen Stimuli mit Ausnahme der Originaläußerungen gepaart werden. Das Vorgehen ergibt insgesamt 14 (9 ungleiche und 5 gleiche) Paare für die erste Stimulusreihe sowie jeweils 15 (9 ungleiche und 6 gleiche) Stimuluspaare für die weiteren Reihen. Alle gepaarten Stimuli werden 5 mal in randomisierter Reihenfolge dargeboten, was 70 (14 x 5) bzw. 75 (15 x 5) Entscheidungen erfordert. Die Anzahl der Wiederholungen wird im Vergleich zum Identifikationstest reduziert, damit die Dauern der beiden Tests vergleichbar lang bleiben und die gesamte Experimentdauer eine Stunde nicht überschreitet.

Die Stimulipräsentation läuft wie folgt ab:

ein 500 ms langer Piepton

750 ms Pause

Stimulus A

2 s Pause

Stimulus X (bzw. A bei physikalisch identischen Kontrollstimuli)

4 s Pause für die Entscheidung

usw. für jedes Stimuluspaar.

Hier werden ebenso Anfangs-, Block- und Endsignale durch Pieptöne gegeben. Der für diese Experimente verwendete Piepton ist ein Sinussignal mit einer Frequenz von 500 Hz, der je nach Funktion unterschiedliche Dauer hat. Der Diskriminationstest dauert für die erste Stimulusreihe 11 min 51 s, für alle weiteren Stimulusreihen je 12 min 43 s.

#### **5.2.4. Versuchspersonen**

An dem Experiment nahmen insgesamt 7 männliche und 31 weibliche Personen aus verschiedenen Städten Rußlands sowie aus anderen GUS-Ländern teil. Das Alter der Versuchspersonen variierte zwischen 20 und 56 Jahren, das mittlere Alter lag bei 27 Jahren. Die Voraussetzung für die Teilnahme an dem Experiment war muttersprachliche Beherrschung der russischen Standardsprache (russ. *literaturnoje proiznoshenije*, vgl. Ozhegov und Shviedova, 1993: 336), was alle Probanden von sich bestätigten. Die Dauer ihres Aufenthalts in Deutschland liegt zwischen 1 und 12 Jahren, wobei die Mehrzahl der Personen einen täglichen Kontakt zur Muttersprache pflegt. Alle Partizipanten hatten keine Vorkenntnisse bezüglich der Fragestellung des Experiments.

#### **5.2.5. Experimentdurchführung**

Die vorliegende Untersuchung umfaßt mehrere Versuchssitzungen, die in einem schallgedämpften Raum am Institut für Phonetik und digitale Sprachverarbeitung Kiel durchgeführt worden sind. Die maximale Anzahl der Teilnehmer pro Sitzung war aus technischen Gründen auf 8 Personen festgelegt. Es fanden aber auch Versuchssitzungen mit weniger als 8 Versuchspersonen statt.

Eine Versuchssitzung bestand aus 3 Tests (s. Kap. 5.2.3.) zu einer der vier hergestellten Stimulusreihen. Die Datenerhebung erfolgte sowohl über Fragebögen als auch über ein Reaktionszeit-Meßgerät RMG3. Die Fragebögen enthielten kurze Aufgabenhinweise. Ausführliche Testinstruktionen wurden mündlich gegeben. Das RMG3-Gerät wurde mit einem Rechner, auf dem die Testsignale gespeichert waren, sowie mit 8 Tastatur-Kästen, über welche die Aufzeichnung lief, verbunden (s. Thon (1982) für einen vergleichbaren Versuchsaufbau). Die Benutzung des RMG3-Geräts erlaubte Messungen und Aufzeichnung der Reaktionszeiten und des Reaktionsverhaltens von allen beteiligten Versuchspersonen sowie die Verarbeitung der erhobenen Daten (wie Derandomisierung, deskriptivstatistische Auswertung).

Am Anfang jeder Testsitzung wurden die Versuchspersonen gebeten, einen kurzen Fragebogen mit persönlichen Daten wie Alter, genaue Herkunft, Fremdsprachenkenntnisse, Kontakt zur Muttersprache und Dauer des Aufenthalts in Deutschland auszufüllen. Die Erhebung von persönlichen Daten war notwendig, um in einer heterogenen Probandengruppe

mögliche Interferenzen mit Zweitsprachen nachvollziehen bzw. ausschließen zu können. In diesem Zusammenhang sei auf Leed (1965) verwiesen, der von intonatorischen Interferenzen bei den im Ausland lebenden Russen berichtet: Diese Personen produzierten in Frageäußerungen steigende statt steigend-fallender Konturen.

Im ersten Test zur seriellen Diskrimination wurden die Versuchspersonen angewiesen, nach dem Abhören einer Serie von Wiederholungen des Satzes „*Jejo zovut Jeliena*“ auf dem Fragebogen zu antworten, ob und wie häufig sie eine Änderung in der Serie wahrnahmen. Nach dem wiederholten Abhören der Serie mußten sie ankreuzen, an welcher Stelle die erste und die eventuellen weiteren, deutlichen Änderungen im Gehörseindruck auftraten. Die letzte Aufgabe des Tests bestand darin, die Bedeutung der ersten Äußerung und der wahrgenommenen Änderungen zu beschreiben. Die Testserie wurde insgesamt dreimal dargeboten.

Im zweiten Test zur Identifikation wurden die Probanden instruiert, die Bedeutung der präsentierten Äußerungen auf einer fünfstufigen Skala, welche die Antwortmöglichkeiten „*eindeutige Aussage*“, „*eher Aussage*“, „*ich weiß nicht*“, „*eher Frage*“, „*eindeutige Frage*“ umfaßte, zu beurteilen und ihre Entscheidung auf dem Fragebogen mit einem Kreuz zu markieren.

Im dritten Test zur AX-Diskrimination wurde die Anweisung gegeben, die paarweise präsentierten Äußerungen auf ihre Gleichheit bzw. Verschiedenheit zu bewerten und die Antwort durch eine einmalige Betätigung des entsprechenden Knopfes auf der Tastatur abzugeben. Die Aufzeichnung des Antwortverhaltens von Versuchspersonen lief über das Reaktionszeit-Meßgerät (s. oben).

Die Signale wurden über einen zentral plazierten Lautsprecher präsentiert. Alle Versuchspersonen saßen in einer ungefähr gleichen Entfernung vom Lautsprecher.

Jede Sitzung dauerte ca. eine Stunde. Im Anschluß äußerten sich alle Probanden dahingehend, daß der Versuch sie ermüdet und zum Ende der Sitzung ihre Aufmerksamkeit nachgelassen hätte. Alle Personen nahmen ausschließlich an einer Testserie teil.

## **5.3. Ergebnisse**

### **5.3.1. Besonderheiten der Stichprobe**

Von den 38 Datenerhebungen wurden lediglich 30 ausgewertet. Die erhobenen Rohdaten sind im Anhang D und E tabellarisch zusammengestellt. Aus Datenschutzgründen werden die Versuchspersonendaten im Folgenden verschlüsselt angegeben. Die Verschlüsselung be-

rücksichtigt das Geschlecht (Kleinbuchstabe), das Namenkürzel (in Großbuchstaben) und das Alter der Probanden.

Eine Versuchsperson (mJMO-56), die an dem Versuch zur zweiten Stimulusreihe teilnahm, perzipierte beim ersten Test zur seriellen Diskrimination keinerlei Änderungen in den vorgeführten Äußerungen und wurde deshalb von den weiteren Tests ausgenommen.

Eine Reihe von Versuchspersonen, die an den Experimenten zu verschiedenen Stimulusreihen teilnahmen (wJPI-21, wVAS-24, wNSI-26, wMTU-24, wMZA-23, wISC-22), interpretierte die wahrgenommenen Veränderungen beim einfachen seriellen Diskriminationstest nicht als Frage, sondern als Kontrast (auch bei der Darbietung des stilisierten Originals der Frageäußerung). Vier dieser Versuchspersonen kommen aus der Ukraine, eine aus Sibirien und eine aus Kaliningrad. Ihre Daten wurden nicht ausgewertet, da deren Sprachkode von dem zu untersuchenden abzuweichen scheint. Dies ist möglicherweise auf einen Unterschied zwischen einzelnen Varietäten des Russischen zurückzuführen. Die Varietätsunterschiede sollen in dieser Arbeit aber nicht weiter thematisiert werden.

Die Daten der Versuchsperson wAKA-39, die an der Testsitzung zur zweiten Stimulusreihe partizipierte, gingen in die Auswertung ebenfalls nicht ein, da die Person Schwierigkeiten mit dem Versuchsablauf hatte und eine große Anzahl von Testaufgaben unbeantwortet ließ. Die gelegentlich fehlenden Antworten der sonstigen Versuchspersonen wurden für den Identifikationstest durch die Antwort „*ich weiß nicht*“ ergänzt.

### **5.3.2. Versuchsplan und Auswertung**

Das durchgeführte Experiment berücksichtigte drei **unabhängige Variablen**: (1) die Anstiegsneigung mit zwei Ausprägungen *steil* vs. *flach* (Gruppenfaktor mit 2 Stufen); (2) die Höhe des Gipfels mit zwei Ausprägungen *hoch* vs. *tief* (Gruppenfaktor mit 2 Stufen); (3) die zeitliche Positionierung des Gipfels mit 11 verschiedenen Ausprägungen in Relation zum Beginn des Akzentvokals. Da für die vorliegende Untersuchung eine späte zeitliche Positionierung des Gipfels in Relation zum Akzentvokal von Interesse ist (vgl. 5.1.), beschränkt sich die geplante statistische Auswertung ausschließlich auf sechs äquidistante Positionen, welche auf dem hergestellten Kontinuum ab dem Onset des Akzentvokals liegen (d.h. 0; +30; +60; +90; +120; +160 ms). Die zeitliche Positionierung des Gipfels (Meßwiederholungsfaktor) hat demnach 6 Stufen. Die Stufen der Gruppenfaktoren waren zum Teil ungleich besetzt: für den Faktor Neigung  $n_1+n_2=14$  und  $n_3+n_4=16$ ; für den Faktor Höhe

$n_1+n_3=15$  und  $n_2+n_4=15$  ( $n_1=7$ ;  $n_2=7$ ;  $n_3=8$ ;  $n_4=8$ ;  $N=n_1+n_2+n_3+n_4=30$  bei ungleich großen Stichproben).

Die *abhängige Variable* stellt das Urteilsverhalten der Probanden in Bezug auf die *Frage-* vs. *Aussage-*Entscheidung in Identifikationstests und die *gleich-* vs. *verschieden-* Meinungen in Diskriminationstests dar. Die Daten wurden in einer Meßwiederholungsprozedur erhoben: Es erfolgte eine 10-fache Abfrage der Urteile zu einem einzelnen Stimulus bei der Identifikation bzw. eine 5-fache Abfrage zu einem Stimuluspaar bei der Diskrimination. Da die Stimuli in randomisierter Reihenfolge präsentiert wurden, sind die Antworten als voneinander unabhängig zu betrachten. Eine mehrfache Datenerhebung pro Versuchsperson liefert stabilisierte Aussagen über das Meinungsbild der Probanden. Alle Antworten werden als gleichwertig angesehen. Die erhobenen Rohdaten wurden wie folgt aufbereitet:

(1) Bei den Identifikationstests wurden die Skalenwerte 4 (*eher Frage*) und 5 (*eindeutige Frage*) mit der Zahl 1 und die Skalenwerte 1 bis 3 (für *eindeutige Aussage*, *eher Aussage*, *ich weiß nicht*, s. Anhang D) mit der Zahl 0 kodiert. Die zehn Urteile jeder Versuchsperson wurden aufaddiert, so daß jede Versuchsperson für ihr Urteilsverhalten pro Stimulus einen Wert von 0 bis 10 erhielt. Durch die nachträgliche Binarisierung der fünfstufigen Skala wird eine Verringerung der Urteilsstreuung erzielt, welche auf persönliche Präferenzen beim Abgeben von definitiven Entscheidungen zurückgeführt werden kann. Zwar ist dieses Vorgehen mit einem Verlust an Informationen zum wahrgenommenen Eindeutigkeitsgrad der beurteilten Äußerungen verbunden, jedoch kann dadurch deutlicher herausgefiltert werden, ob eine Kategoriengrenze auf dem Gipfelverschiebungskontinuum vorliegt. Für die Fragestellung der vorliegenden Arbeit ist dies von zentralem Interesse.

(2) Bei den Diskriminationstests wurde in ähnlicher Weise vorgegangen: Die *verschieden-*Urteile (Skalenwert 2 bei der Datenerhebung, s. Anhang E) wurden mit der Zahl 1, die *gleich-*Urteile (Skalenwert 1, Anhang E) sowie fehlende Antworten mit der Zahl 0 kodiert. Hier konnte das Entscheidungsverhalten eines jeden Probanden durch einen Wert von 0 bis 5 wiedergegeben werden.

Die Bedingungen der durchgeführten Experimente ergeben einen Versuchsplan mit drei Faktoren (s. Tabellen 5./4. und 5./5.).

Tab. 5./4.: Tabellarische Darstellung des Versuchsplans für die Identifikationsexperimente.

Gruppenfaktoren		Meßwiederholungsfaktor					
		0	+30	+60	+90	+120	+150
flach	tief	n <sub>1a</sub>	n <sub>1b</sub>	...	...	...	n <sub>1f</sub>
flach	hoch	n <sub>2a</sub>	...	...	...	...	...
steil	tief	n <sub>3a</sub>	...	...	...	...	...
steil	hoch	n <sub>4a</sub>	...	...	...	...	...

Tab. 5./5.: Tabellarische Darstellung des Versuchsplans für die Diskriminationsexperimente.

Gruppenfaktoren		Meßwiederholungsfaktor					
		-60/ 0	-30/ +30	0/ +60	+30/ +90	+60/ +120	+90/ +150
flach	tief	n <sub>1a</sub>	n <sub>1b</sub>	...	...	...	n <sub>1f</sub>
flach	hoch	n <sub>2a</sub>	...	...	...	...	...
steil	tief	n <sub>3a</sub>	...	...	...	...	...
steil	hoch	n <sub>4a</sub>	...	...	...	...	...

Dieser Aufbau entspricht für beide Testarten (Diskrimination und Identifikation) einem 22.6-Plan (*split plot factorial design*) nach Kirk (1968). Ein für den vorliegenden Versuchsplan geeignetes statistisches Verfahren stellt die multifaktorielle univariate Varianzanalyse mit Meßwiederholungen dar. Diese prüfstatische Methode ermöglicht eine simultane Kontrolle mehrerer Variablen und deren Interaktionen in ihrer Auswirkung auf die Unterschiede im Verhalten von Versuchspersonen (vgl. Bortz, 1979: 295ff.).

Die varianzanalytische Voraussetzung der Intervallskalierung der abhängigen Variable wird als erfüllt angenommen. Die weiteren Voraussetzungen betreffen die Normalverteilung und Varianzhomogenität der Fehleranteile der abhängigen Variable. Von einer Normalverteilung der Fehlerkomponenten wird ohne Überprüfung ausgegangen (vgl. Bortz, 1979: 344). Die Varianzhomogenität überprüft der *Levene-Test* (mit  $p = 0,20$ ; Bortz, 1979: 346). Das Alpha-Kriterium wird auf  $p = 0,01$  gesetzt. Die Rechnung des oben erläuterten prüfstatischen Verfahrens (ANOVA) wurde im Psychologischen Institut der CAU Kiel mit dem dort verfügbaren Programmpaket SPSS (Brosius, 2002) durchgeführt. Die für die Varianzanalyse vorbereiteten Datensätze befinden sich im Anhang H.

Auf eine inferenzstatistische Überprüfung der vierten Hypothese (s. Abschnitt 5.1.) wird hier verzichtet. Sie soll im Kapitel 6 anhand der deskriptiven Statistiken diskutiert werden.

Für beide Testserien (Identifikation und Diskrimination) wurden außerdem deskriptive Übersichten anhand der beobachteten Häufigkeitsverteilung erstellt. Für die Reihe der Identifikationstests wurde eine Binarisierung der Verteilung nach zwei Kategorien *Frage* vs. *Aussage* vorgenommen. Hierbei wurden ausschließlich die Antworten *eher Frage* (Skalenwert 4 bei der Datenerhebung, vgl. Anhang D) und *eindeutige Frage* (Skalenwert 5) zur Kategorie *Frage* zusammengefaßt, da im Mittelpunkt des durchgeführten Experiments die intonatorische Kodierung der Frageintention steht und in dieser Kategorie keine ambigen Stimuli berücksichtigt werden sollen (vgl. Darstellung der abhängigen Variable oben). Die Antworten *eindeutige Aussage*, *eher Aussage* und *ich weiß nicht* (Skalenwerte 1, 2 und 3 bei der Datenerhebung, vgl. Anhang D) wurden folgerichtig der Kategorie *Aussage* zugeordnet. Das Resultat der vorgenommenen Zusammenfassung der Rohdaten ist in den Tabellen des Anhangs F dargestellt.

Die deskriptiven Statistiken für die Rohdaten der durchgeführten Diskriminationstests sind dem Anhang G zu entnehmen. Dort werden die absoluten und relativen Häufigkeiten der *gleich-* vs. *verschieden-*Urteile sowie der fehlenden Antworten getrennt nach Gipfelgestalt und Art der Stimuluspaarbildung (physikalisch gleiche vs. ungleiche Paare) präsentiert. Die bei der Testdurchführung über das RMG 3 gemessenen Reaktionszeiten der Versuchspersonen werden in der Datenauswertung vernachlässigt.

Abschließend sei angemerkt, daß sich die Angaben zu den relativen Häufigkeiten (in %) bei den deskriptivstatistischen Übersichten in die individuellen Meßwerte der inferenzstatistischen Auswertung (und umgekehrt) umdeuten lassen, da sie auf dem gleichen Auswertungsprinzip beruhen (s. oben).

### **5.3.3. Ergebnisse der Identifikationstests**

Im folgenden Abschnitt werden die Ergebnisse des Identifikationstests für die vier getesteten Gipfelgestalten in Form von Tabellen und Graphiken aufgeführt, wobei sowohl eine deskriptive Übersicht gegeben als auch die Prüfstatistik berücksichtigt wird.

Zunächst werden die Zusammenhänge zwischen den spezifischen Gipfelprägungen und den Differenzen im Entscheidungsverhalten der Probanden (Anhang F) graphisch verdeutlicht. Die relativen Häufigkeiten der extrahierten *Frage*-Urteile sind als Identifikationsfunktionen über dem zeitlichen Kontinuum der unterschiedlichen Gipfelpositionen in Relation zum Onset des Akzentvokals für die vier getesteten Gipfelgestalten in Graphik 5./9.

abgebildet. Hierbei wird die Bewertung der stilisierten Originale der Frage- und Aussage-Äußerungen berücksichtigt.

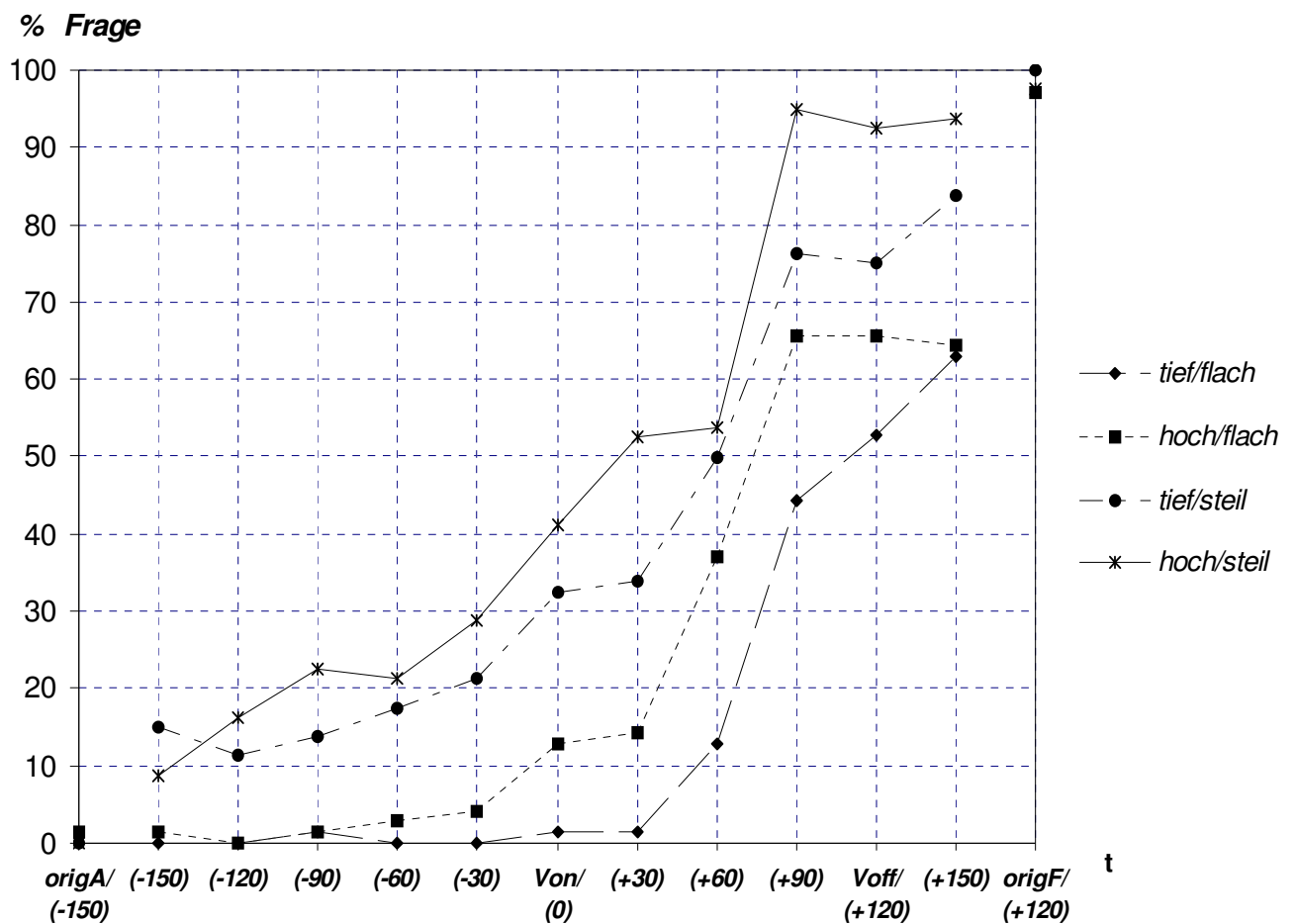


Abb. 5./9.: Identifikationsfunktionen der vier getesteten Gipfelgestalten *tief/flach* ( $n_1=7$ ) vs. *hoch/flach* ( $n_2=7$ ) vs. *tief/steil* ( $n_3=8$ ) vs. *hoch/steil* ( $n_4=8$ ) im Vergleich.

Anhand der präsentierten Identifikationsfunktionen können folgende Tendenzen festgestellt werden:

(1) Die Stimuli der zweiten Hälfte des Kontinuums (Positionierung des Gipfels nach dem Vokalonset) rufen unabhängig von der jeweiligen Gipfelgestalt mit zunehmender Häufigkeit *Frage*-Urteile hervor.

(2) Die 50%-Grenze findet sich an unterschiedlichen Stellen des zeitlichen Kontinuums in Abhängigkeit von der jeweiligen Gipfelgestalt. So zeigt die Verteilung für die Gipfelkontur *hoch/steil* bereits zwischen dem Vokalonset und dem ersten Drittel des Akzentvokals die 50%-Grenze. Bei der Gipfelgestalt *tief/steil* liegt sie in der Mitte des

Akzentvokals (bei +60 ms). Viel später im Kontinuum wird die Grenze bei der Gipfelgestalt *hoch/flach* (zwischen +60 und +90 ms) und *tief/flach* (zwischen +90 ms und Vokaloffset) erreicht.

(3) In allen Häufigkeitsverteilungen liegt ein bedeutsamer Anstieg der *Frage*-Urteile beim Wechsel der Gipfelposition von der Mitte (+60 ms) in das letzte Drittel (+90 ms) des Akzentvokals vor. Dieser Übergang und der dabei erreichte Prozentsatz an *Frage*-Entscheidungen ist unterschiedlich stark ausgeprägt: Die Gipfelgestalt *hoch/steil* ruft den deutlichsten Anstieg von ca. 40% hervor, wobei 95% der Äußerungen als Fragen identifiziert werden. Etwas weniger ausgeprägt zeigt sich die ca. 25%-ige Urteilserhöhung bis 76% für die getestete Gipfelgestalt *tief/steil* und bis ca. 66% für die Gipfelgestalt *hoch/flach*. Die Zunahme von ca. 30% ist für die Gipfelkontur *tief/flach* zu registrieren, wobei die Identifikationsrate von 45% am schlechtesten ausfällt.

(4) Für die Gipfelgestalten *hoch/steil* und *tief/steil* ist bereits im Bereich der frühen Gipfelpositionen (von -150 ms bis Vokaloffset) ein Anstieg der *Frage*-Urteile im Vergleich zu den Konturen *hoch/flach* und *tief/flach* zu vermerken.

(5) Die stilisierten Originale der natürlich produzierten *Frage*- und Aussage-Äußerungen, die als Kontrollstimuli in den Identifikationstests fungierten, werden stets mit beinahe 100%-iger Übereinstimmung von den Versuchspersonen als solche identifiziert.

Das Hauptergebnis der oben dargestellten deskriptiven Statistik zu den Identifikationstests ist, daß ab der Gipfelpositionierung spät im Akzentvokal (+90 ms) für alle getesteten Gipfelgestalten eine deutliche Zunahme an *Frage*-Urteilen zu vermerken ist. Für die späten Gipfel mit der Gestalt *hoch/steil* liegt die höchste Urteilsrate von über 90% vor.

Die folgenden Tabellen 5./6.-7. präsentieren die inferenzstatistischen Ergebnisse der durchgeführten univariaten mehrfaktoriellen Varianzanalyse mit Meßwiederholungen für die Identifikation. Bei den Gruppenfaktoren zeigt der Faktor Neigung einen signifikanten Einfluß auf die untersuchte abhängige Variable (*Frage*-Urteile). Die Kombination der Gruppenfaktoren mit dem Meßwiederholungsfaktor liefert zwei signifikante Befunde: Über die Gesamtheit der getesteten Gipfelgestalten sind die Urteilsänderungen bei der Gipfelpositionsverschiebung (1) von +30 zu +60 sowie (2) von +60 zu +90 signifikant. Sonst liegen keine weiteren Befunde vor.

Tab. 5./6.: Ergebnisse der multifaktoriellen univariaten Varianzanalyse für die Identifikation: Haupt- und Interaktionseffekte unter den Gruppenfaktoren (Zwischensubjekteffekte).

<i>Gruppenfaktoren</i>	<i>Ausprägung</i>	<i>F</i>	<i>Signifikanz</i>
Neigung	<i>(flach/tief, flach/hoch) vs. (steil/tief, steil/hoch)</i>	15,975	,000*
Höhe	<i>(hoch/flach, hoch/steil) vs. (tief/flach, tief/steil)</i>	3,361	,078
Neigung und Höhe	<i>flach/tief vs. flach/hoch vs. steil/tief vs. steil/hoch</i>	,000	,997

Tab. 5./7.: Einzelvergleiche von Stufe zu Stufe auf dem Meßwiederholungsfaktor (wiederholte Kontraste) für die Identifikation.

<i>Gruppenfaktoren</i>	<i>Meßwiederholungsfaktor</i>	<i>F</i>	<i>Signifikanz</i>
Gesamtbefund ohne Spezifizierung der Gruppenfaktoren	0 zu +30	,851	,365
	+30 zu +60	8,852	,006*
	+60 zu +90	56,169	,000*
	+90 zu +120	,897	,352
	+120 zu +150	3,763	,063
Neigung: <i>(flach/tief, flach/hoch) vs. (steil/tief, steil/hoch)</i>	0 zu +30	,538	,470
	+30 zu +60	,930	,344
	+60 zu +90	,194	,663
	+90 zu +120	2,627	,117
	+120 zu +150	,602	,445
Höhe: <i>(hoch/flach, hoch/steil) vs. (tief/flach, tief/steil)</i>	0 zu +30	,573	,456
	+30 zu +60	,042	,839
	+60 zu +90	,509	,482
	+90 zu +120	1,949	,174
	+120 zu +150	3,954	,057
Wechselwirkung Höhe und Neigung: <i>flach/tief vs. flach/hoch vs. steil/tief vs. steil/hoch</i>	0 zu +30	,322	,575
	+30 zu +60	2,305	,141
	+60 zu +90	1,102	,304
	+90 zu +120	1,373	,252
	+120 zu +150	,002	,962

Wie die Ergebnisse des *Levene-Tests* (s. Tabelle 5./8.) zeigen, sind die varianzanalytischen Voraussetzungen in den Daten zur Identifikation häufig verletzt. Diese Verletzungen werden in erster Linie dadurch relativiert, daß ein konservatives Alpha-Niveau gewählt wurde. Darüber hinaus wurde durch mehrere Arbeiten belegt, daß sich die Varianzanalyse robust gegen die Voraussetzungsverletzungen verhält und auch dann zu richtigen Entscheidungen führt (vgl. Bortz, 1979: 347). In diesem Zusammenhang kann festgehalten werden, daß sich trotz der verhältnismäßig geringen Zahl an Versuchspersonen eine klare Befundstruktur zeigt. Die gewonnenen Ergebnisse der Varianzanalyse stehen im Einklang mit der deskriptiven Datenbetrachtung und sind im vorliegenden Fall interpretierbar.

Tab. 5./8.: Ergebnisse des *Levene-Tests* für die Identifikation.

	<i>Meßwiederholungsfaktor</i>					
	0	+30	+60	+90	+120	+150
F	14,591	10,735	,910	4,773	2,879	2,428
Signifikanz	,000*	,000*	,450	,009*	,055*	,088*

### **5.3.4. Ergebnisse der Diskriminationstests**

Im folgenden Abschnitt werden die deskriptiven sowie prüfstatistischen Ergebnisse der durchgeführten Diskriminationstests für die vier getesteten Gipfelgestalten dargestellt.

Die beobachteten (relativen) Häufigkeiten der *verschieden*-Urteile (Anhang G) werden in den Vergleichsgraphiken als Diskriminationsfunktionen getrennt nach Art der Stimuluspaarbildung (physikalisch gleiche vs. ungleiche Paare) dargestellt (Abb. 5./10.-11.). Auf der Zeitachse sind die in den Tests realisierten Positionen des F0-Gipfels in Relation zum Beginn des Akzentvokals (0) markiert. Die relativen Häufigkeiten der *verschieden*-Urteile (vertikale Achse) werden hierbei mit Bezug auf den zweiten Stimulus eines Paares abgebildet. Die physikalisch verschiedenen Stimuluspaare weisen einen zeitlichen Abstand von 60 ms zwischen den einzelnen Gipfelpositionierungen auf.

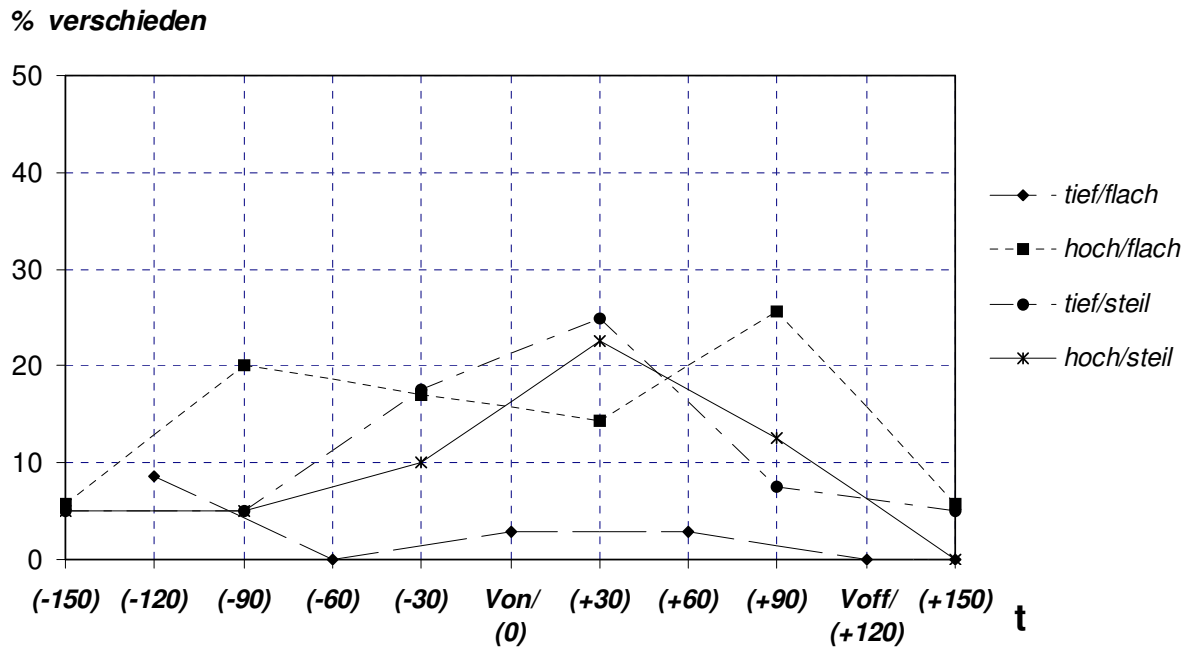


Abb. 5./10.: Experimentell gemessene Diskriminationsleistungen für physikalisch identische Paare aus den gestesteten Kontinua der vier Gipfelgestalten *tief/flach* ( $n_1=7$ ) vs. *hoch/flach* ( $n_2=7$ ) vs. *tief/steil* ( $n_3=8$ ) vs. *hoch/steil* ( $n_4=8$ ) im Vergleich.

Der Vergleich zwischen den Diskriminationskurven, die für physikalisch identische Paare gewonnen wurden (Abb. 5./10.), zeigt, daß alle Paare aus der ersten Stimulusreihe (Gipfelkontur mit der Gestalt *tief/flach*) ein klares *gleich*-Urteil erzeugten. Die meisten *verschieden*-Urteile für die getesteten Paare wurden bei den Stimuli der zweiten Stimulusreihe (Gipfelgestalt *hoch/flach*) registriert. Die Stimuluspaare der dritten und der vierten Stimulusreihe (Gipfelkonturen *tief/steil* und *hoch/steil*) führten zu einem ähnlichen Urteil, wobei neben den überwiegend als identisch bewerteten Zweiergruppen ein Diskriminationsmaximum für die Stimuli mit der Gipfelpositionierung im ersten Drittel des Akzentvokals (+30 ms) deutlich wird.

**% verschieden**

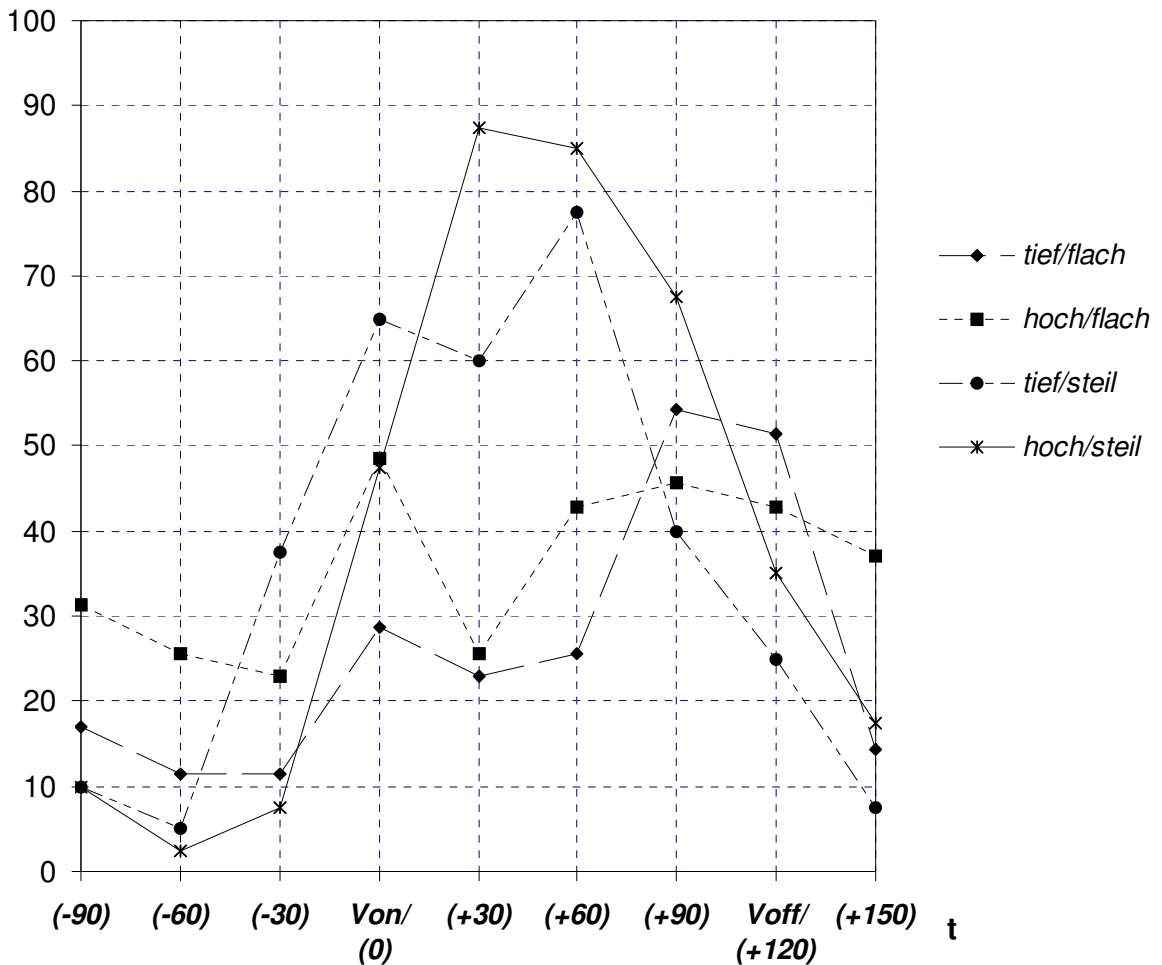


Abb. 5./11.: Experimentell gemessene Diskriminationsleistungen für physikalisch verschiedene Paare aus den gestesteten Kontinua der vier Gipfelgestalten *tief/flach* ( $n_1=7$ ) vs. *hoch/flach* ( $n_2=7$ ) vs. *tief/steil* ( $n_3=8$ ) vs. *hoch/steil* ( $n_4=8$ ) im Vergleich.

Die Diskrimination von physikalisch ungleichen Paaren zeigt die allgemeine Tendenz, daß die Stimuluspaare in der Mitte des Kontinuums besser diskriminiert werden als die gepaarten Randstimuli (vgl. Abb. 5./11.). Das Urteilsverhalten der Probanden fällt in Abhängigkeit von der jeweiligen getesteten Gipfelgestalt allerdings unterschiedlich aus. Die festgestellten Unterschiede lassen sich wie folgt zusammenfassen:

(1) Die Bildung eines Diskriminationsmaximums erfolgt an verschiedenen Stellen entlang des Kontinuums der äquidistant geänderten zeitlichen Gipfelsynchronisierung: Hierbei liegt für die hohe, steil ansteigende Gipfelgestalt der vierten Stimulusreihe eine Diskriminationsgrenze zwischen einer frühen und einer mittleren Positionierung. Die gleiche

Grenze ist für die dritte Stimulusreihe (Gipfelkontur *tief/steil*) festzuhalten. In der erste Stimulusreihe (Gipfelkontur *tief/flach*) tritt eine Diskriminationsgrenze zwischen einer mittleren und einer späten Gipfelposition auf. Die zweite Stimulusreihe (Gipfelgestalt *hoch/flach*) zeigt zwei Diskriminationsmaxima, die zum einen zwischen einer extrem frühen (vor dem Vokalonset) und einer frühen (am Vokalonset) Gipfelpositionierung und zum zweiten zwischen einer eher mittleren und einer eher späten zeitlichen Gipfelsynchronisierung lokalisierbar sind. Das zweite Diskriminationsmaximum ist uneindeutig.

(2) Die Ausprägung der Maximumbildung kommt mit ca. 88% für die vierte Stimulusreihe (Gipfelgestalt *hoch/steil*) am deutlichsten zum Vorschein. Vergleichbar fällt das Ergebnis (ca. 78%) für die dritte Stimulusreihe aus. Weniger stark ist die 54%-ige Maximumbildung in der ersten Stimulusreihe (Gipfel *tief/flach*). Der Verlauf der Diskriminationsfunktion für die zweite Stimulusreihe (Gipfel *hoch/flach*) überschreitet an keiner Stelle die 50%-Grenze.

(3) Die Stimulusreihen mit den Gipfelgestalten *tief/flach*, *hoch/flach* und *tief/steil* deuten ein zusätzliches, weniger ausgeprägtes Diskriminationsmaximum an der Stelle zwischen einer extrem frühen und frühen Gipfelposition an. Diese Beobachtung steht im Gegensatz zur Reihe mit der getesteten Kontur *hoch/steil*, die ausschließlich eine Maximumbildung aufweist.

Das Hauptergebnis der oben detailliert dargestellten deskriptiven Statistiken liegt darin, daß bei allen Gipfelgestalten die physikalisch ungleichen Stimuluspaare in der Mitte des Kontinuums besser diskriminiert werden als die gepaarten Randstimuli. Die physikalisch identischen Stimuluspaare wurden überwiegend als gleich bewertet.

Die Ergebnisse der univariaten mehrfaktoriellen Varianzanalyse für die Diskrimination werden in den Tabellen 5./9.-10. dargestellt. Bei den Gruppenfaktoren können keine signifikanten Effekte auf dem gesetzten Alpha-Niveau festgestellt werden. Es besteht aber ein signifikanter Trend für den Faktor Neigung. In der Kombination der Gruppenfaktoren mit dem Meßwiederholungsfaktor werden insgesamt fünf signifikante Befunde deutlich. Über die Gesamtheit der getesteten Gipfelgestalten ist der Unterschied in der Bewertung der Stimuluspaare mit Gipfelpositionierungen (1) +30/ +90 vs. +60/ +120 und (2) +60/ +120 vs. +90/ +150 signifikant. Zwei weitere Befunde liegen für den Neigungsfaktor vor: Das Urteil der Probanden über die Stimuluspaare (3) -60/ 0 vs. -30/+30 und (4) 0/ +60 vs. +30/ +90 wird in Abhängigkeit von der Gipfelgestalt mit einer steilen vs. flachen Anstiegsneigung signifikant beeinflusst. Der letzte signifikante Befund betrifft die Bewertung der Stimuluspaare (5) -60/ 0

vs. -30/+30 bei der Wechselwirkung zwischen allen Bedingungen unter den Gruppenfaktoren.

Tab. 5./9.: Ergebnisse der multifaktoriellen univariaten Varianzanalyse für die Diskrimination: Haupt- und Interaktionseffekte unter den Gruppenfaktoren (Zwischensubjekteffekte).

<i>Gruppenfaktoren</i>	<i>Ausprägung</i>	<i>F</i>	<i>Signifikanz</i>
Neigung	<i>(flach/tief, flach/hoch) vs. (steil/tief, steil/hoch)</i>	4,708	,039
Höhe	<i>(hoch/flach, hoch/steil) vs. (tief/flach, tief/steil)</i>	1,884	,182
Neigung und Höhe	<i>flach/tief vs. flach/hoch vs. steil/tief vs. steil/hoch</i>	,057	,813

Tab. 5./10.: Einzelvergleiche von Stufe zu Stufe auf dem Meßwiederholungsfaktor (wiederholte Kontraste) für die Diskrimination.

<i>Gruppenfaktoren</i>	<i>Meßwiederholungsfaktor</i>	<i>F</i>	<i>Signifikanz</i>
Gesamtbefund ohne Spezifizierung der Gruppenfaktoren	-60/ 0 vs. -30/ +30	,110	,734
	-30/ +30 vs. 0/ +60	2,888	,101
	0/ +60 vs. +30/ +90	1,082	,308
	+30/ +90 vs. +60/ +120	7,620	,010*
	+60/ +120 vs. +90/ +150	10,810	,003*
Neigung: <i>(flach/tief, flach/hoch) vs.</i> <i>(steil/tief, steil/hoch)</i>	-60/ 0 vs. -30/ +30	10,710	,003*
	-30/ +30 vs. 0/ +60	,059	,810
	0/ +60 vs. +30/ +90	14,543	,001*
	+30/ +90 vs. +60/ +120	4,699	,040
	+60/ +120 vs. +90/ +150	,110	,743
Höhe: <i>(hoch/flach, hoch/steil) vs.</i> <i>(tief/flach, tief/steil)</i>	-60/ 0 vs. -30/ +30	2,056	,163
	-30/ +30 vs. 0/ +60	,077	,784
	0/ +60 vs. +30/ +90	,064	,803
	+30/ +90 vs. +60/ +120	,824	,372
	+60/ +120 vs. +90/ +150	1,762	,196
Wechselwirkung zwischen Höhe und Neigung: <i>flach/tief vs. flach/hoch vs.</i> <i>steil/tief vs. steil/hoch</i>	-60/ 0 vs. -30/ +30	10,234	,004*
	-30/ +30 vs. 0/ +60	2,771	,108
	0/ +60 vs. +30/ +90	4,069	,054
	+30/ +90 vs. +60/ +120	,824	,372
	+60/ +120 vs. +90/ +150	1,762	,196

Wie aus der folgenden Tabelle 5./11. ersichtlich, ist die varianzanalytische Voraussetzung in den vorliegenden Diskriminationsdaten in den meisten Fällen erfüllt. Hier gelten ebenfalls die Überlegungen, die zu den Ergebnissen des *Levene-Tests* für die Identifikation angeführt wurden (s. 5.3.3.).

Tab. 5./11.: Ergebnisse des *Levene-Tests* für die Diskrimination.

	<i>Meßwiederholungsfaktor</i>					
	-60/ 0	-30/ +30	0/ +60	+30/ +90	+60/ +120	+90/ +150
F	,436	2,193	1,094	1,241	1,594	1,026
Signifikanz	,729	,113*	,369	,315	,215	,397

### 5.3.5. Ergebnisse der Vortests

In diesem Abschnitt werden die Ergebnisse des einfachen seriellen Diskriminationstests zusammengefaßt, der am Anfang der Versuchssitzungen als Vortest durchgeführt wurde. Die zeitliche Position des Maximums der jeweiligen Gipfelkontur ist stets in Relation zum Onset des Akzentvokals angegeben. Für die erste Stimulusreihe, in der die Gipfelkontur mit der Gestalt *tief/flat* getestet wurde, ergab sich folgende Häufigkeitsverteilung für die Wahrnehmung von Änderungen bei einer Gruppe von 7 Versuchspersonen (s. Tab. 5./12.):

Tabelle 5./12.: Absolute Häufigkeiten der wahrgenommenen Änderungen in der ersten Stimulusreihe beim einfachen seriellen Diskriminationstest in der Gruppe von 7 Versuchspersonen.

<i>Wahrnehmung von Änderungen</i>	<i>Position der Gipfelkontur „tief/flat“</i>					
	+30	+60	+90	+120	+150	OrigF
erste Änderung	1	-	1	2	-	3
weitere Änderungen	-	-	-	-	-	4
<b>Gesamt</b>	<b>1</b>	<b>-</b>	<b>1</b>	<b>2</b>	<b>-</b>	<b>7</b>

Die erste Änderung wurde an unterschiedlichen Stellen in der Reihe wahrgenommen: (1) beim Übergang von einer frühen zu einer mittleren Gipfelposition (+30 ms nach dem Vokaloffset) und (2) beim Übergang von einer mittleren zu einer eher späten Gipfelposition (im letzten Drittel des Akzentvokals (+90 ms) und am Vokaloffset). Diese Änderungen

wurden in Verbindung mit *Zweifel*, *Unsicherheit* oder *Ungewißheit* gebracht. Das stilisierte Original der Frageäußerung wurde von allen Versuchspersonen als anders im Vergleich zu vorangehenden Stimuli gehört und als *eine eindeutige Frage* verstanden. Drei der beteiligten Versuchspersonen nahmen lediglich einen Unterschied wahr, nämlich den zwischen dem Frage-Original und den übrigen Stimuli. Das stilisierte Original der Aussage wurde von allen Versuchspersonen übereinstimmend als *Behauptung*, *Auskunft* gedeutet.

Die Ergebnisse des Vortests der zweiten Stimulusreihe, für welche die Gipfelkontur *hoch/flach* resynthetisiert wurde, sind in der Tabelle 5./13. präsentiert.

Tabelle 5./13.: Absolute Häufigkeiten der wahrgenommenen Änderungen in der zweiten Stimulusreihe beim einfachen seriellen Diskriminationstest in einer Gruppe von 7 Versuchspersonen.

Wahrnehmung von Änderungen	Position des Gipfels mit der Gestalt „hoch/flach“												
	-150	-120	-90	-60	-30	Von	+30	+60	+90	+120	+150	Orig Voff	F
erste Änderung	2	1	-	-	-	1	-	-	-	-	1	2	
weitere Änderungen	-	-	-	-	-	-	1	1	-	2	-	3	
<b>Gesamt</b>	<b>2</b>	<b>1</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>-</b>	<b>2</b>	<b>1</b>	<b>5</b>	

Die Bedeutung der originalen stilisierten Aussageäußerung wurde im zweiten Experiment als *Behauptung*, *neutrale Auskunft*, *Feststellung*, *Information*, *Bestätigung der Information*, *feste Sicherheit* interpretiert. Die erste Änderung in der vorgeführten Reihe wurde zwischen der Testkontur und dem Aussageoriginal mit gleicher zeitlicher Gipfelpositionierung (bei -150 ms) perzipiert, wobei die Testkontur mit einem höheren Gipfel den Eindruck der *Hervorhebung des Namens* erzeugte. Andere Versuchspersonen nahmen die erste Änderung in den nachfolgenden Stimuli (mit dem Gipfelmaximum bei -120 ms und am Vokalonset) wahr, ihnen kam jedoch dieselbe Interpretation zu. Andererseits wurde die erste Änderung erst spät in der Stimulusreihe (bei der Position des Gipfelmaximums bei +150 ms) gehört und semantisch mit *Unsicherheit* (des Sprechers) in Verbindung gebracht. Weitere Änderungen wurden in der Mitte des Akzentvokals und am Vokaloffset vermerkt und als *weniger behauptend* bzw. *fragend* interpretiert. Zwei Personen nahmen lediglich die Unterscheidung zwischen dem Frage-Original und den restlichen Äußerungen wahr. Die stilisierte Originalkontur der Frage erzeugte bei allen 7 Versuchspersonen das Perzept einer *eindeutigen Frage*.

Die Häufigkeitstabelle 5./14. zum Vortest der dritten Stimulusreihe gibt den Überblick über das Verhalten der Versuchspersonen, die mit der Positionierungsvariation der Gipfelgestalt *tief/steil* auf der Zeitachse konfrontiert wurden.

Tabelle 5./14.: Absolute Häufigkeiten der wahrgenommenen Änderungen in der dritten Stimulusreihe beim einfachen seriellen Diskriminationstest in einer Gruppe von 8 Versuchspersonen.

Wahrnehmung von Änderungen	Position des Gipfels mit der Gestalt „tief/steil“												
	-150	-120	-90	-60	-30	Von	+30	+60	+90	+120	+150	Orig Voff	F
erste Änderung	1	2	1	1	2	1	-	-	-	-	-	-	1
weitere Änderungen	-	-	-	-	-	-	2	-	-	1	1	-	7
<b>Gesamt</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>2</b>	-	-	<b>1</b>	<b>1</b>	-	<b>8</b>

Die einzelnen Versuchspersonen nahmen die erste Änderung bei unterschiedlichen Gipfelpositionen der dargebotenen Kontur *tief/steil* wahr. Die Variation befindet sich jedoch im Rahmen eines frühen Gipfels (mit dem Maximum vor dem Akzentvokal bzw. am Vokalonset). Die Stimuli aus diesem Bereich wurden semantisch als *Ausruf, Hervorhebung des Namens, Verstärkung der Betonung, Kontrast zum anderen Namen* eingestuft und auch bei der Identifikation als *etwas gereizte Reaktion auf unkorrekte Namensnennung* in Verbindung mit Emotionen gebracht. Eine weitere Änderung, die beim Übergang von einer frühen zu einer mittleren Gipfelposition mit dem Maximum bei +30 ms wahrgenommen wurde, blieb ohne semantische Interpretation. Die wahrgenommenen Änderungen am Vokaloffset (bei +120 und +150 ms) wurden als *Zweifel* oder *Frage* interpretiert. Das stilisierte Original der Aussage interpretierten die Versuchspersonen als *eine Behauptung, Bestätigung, neutrale Information* oder auch *Aussage(satz)*. Bloß eine Versuchsperson nahm eine einzige Unterscheidung zwischen dem Kontrollstimulus des Frageoriginals und den restlichen Äußerungen in der Serie vor. Die Probanden waren sich einig, daß es sich beim stilisierten Original der Frageäußerung tatsächlich um eine *Frage* handelt.

In der Tabelle 5./15. sind die Vortestergebnisse für die vierte Stimulusreihe dargestellt, in der die Auswirkungen der unterschiedlichen zeitlichen Positionierungen der Gipfelgestalt *hoch/steil* auf die Wahrnehmung zu testen waren.

Tabelle 5./15.: Absolute Häufigkeiten der wahrgenommenen Änderungen in der vierten Stimulusreihe beim einfachen seriellen Diskriminationstest in der Gruppe von 8 Versuchspersonen.

Wahrnehmung von Änderungen	Position des Gipfels mit der Gestalt „hoch/steil“											
	-150	-120	-90	-60	-30	Von	+30	+60	+90	+120	+150	Orig Voff F
erste Änderung	2	-	-	-	-	2	5	-	-	-	-	-
weitere Änderungen	-	-	-	-	-	-	2	-	3	-	-	4
<b>Gesamt</b>	<b>2</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>2</b>	<b>7</b>	<b>-</b>	<b>3</b>	<b>-</b>	<b>-</b>	<b>4</b>

Als erstes wurde ein Unterschied zwischen dem Original der Aussage und der manipulierten F0-Kontur mit gleicher zeitlicher Gipfelpositionierung (bei -150 ms) vermerkt. Das Original wird hierbei als *eine Mitteilung, neutrale Information, Vorstellung einer Person (am Anfang eines Dialogs), Behauptung*, die Änderung dagegen als *Kontrast* oder *Vermutung* interpretiert. Viele andere Versuchspersonen perzipierten erstmals eine Änderung in der vorgeführten Reihe bei den Äußerungen mit dem Gipfelmaximum am Vokalonset und im ersten Drittel des Vokals. Bei der semantischen Interpretation des wahrgenommenen Unterschieds spalteten sich die Meinungen sehr stark: Einerseits klangen diese Äußerungen für die Probanden *weniger behauptend* oder als *Frage, Nachfrage*, andererseits wurden sie als *eine kontrastierende Hervorhebung des Namens* interpretiert. Bei einer weiteren Änderung, die später in der Reihe wahrgenommen wurde und der Äußerung mit dem Gipfelmaximum im letzten Drittel des Akzentvokals entspricht, waren sich die Versuchspersonen einig, daß es sich um eine Frage handelte. Einige Personen bemerkten den Unterschied des stilisierten Originals der Frage zum Rest der Reihe. Diese Interpretation war jedoch nicht semantischer, sondern phonetischer Natur, da sie die Lautstärke und die Stimmqualität im stilisierten Frage-Original betraf: Das Original kam den betreffenden Personen als *lauter (als ob in einer lauten Umgebung gefragt wird), heller und deutlicher gesprochen* vor. Bloß eine Versuchsperson nahm in diesem Experiment erst beim Hören des Originals eine Frage wahr. Alle Versuchspersonen registrierten mehr als eine Änderung in der vorgeführten Reihe.

Zusammenfassend läßt sich zu den Ergebnissen der oben beschriebenen Vortests sagen, daß die Konturen mit einer frühen Gipfelposition als Aussagen (oder *Behauptungen*), die mit einem späten Gipfel als Fragen (oder auch *Unsicherheit, Zweifel* des Sprechers) semantisch eingeordnet werden. Für die resynthetisierten Gipfelgestalten *hoch/flach, tief/steil* und *hoch/steil* wurde außer der Differenzierung zwischen Frage und Aussage eine weitere semantische Kategorisierung registriert, nämlich der Kontrast (oder *verstärkte Hervorhebung*

*des Namens*). Die Äußerung mit einem mittleren Gipfel der Kontur *hoch/steil* wurde von einigen Versuchspersonen als Frage, von anderen jedoch als Kontrast beurteilt.

## 6. Diskussion

### 6.1. Interpretation der Ergebnisse

Ein im traditionellen Sinne kategorialer Wahrnehmungswechsel, der auf einer Korrelation zwischen einem Diskriminationsmaximum und einer starken Identifikationsänderung beruht (Repp, 1984; Unterkapitel 2.1.5.), kann anhand der Ergebnisse deskriptiver Statistiken für die getesteten Gipfelgestalten *flach/tief*, *steil/tief* und *steil/hoch* festgestellt werden. Die Ergebnisse für die Gipfelgestalt *flach/hoch* zeigen in der Identifikation zwar einen Übergang der Urteile von der Aussage in die Frage-Kategorie bei einer späten Gipfelpositionierung, in der Diskrimination bleibt jedoch eine entsprechende ausgeprägte Maximumbildung aus (vgl. Abb.5./9. und 5.11).

Für alle getesteten Gipfelgestalten kann ein Kategoriensprung von Aussage zu Frage bei der Änderung der Gipfelposition von mittel zu spät (+60 zu +90 ms) registriert werden. Hierbei ist zu vermerken, daß bei der späten Gipfelposition (+90 ms) durch die Präsenz jedes zusätzlichen, die Gipfelgestalt betreffenden Parameters (vgl. Abschnitt 5.2.2.2.) die *Frage*-Identifikation verbessert wird. Ähnliche Tendenzen kristallisieren sich bei der Diskrimination heraus: Abgesehen von der Kontur *hoch/flach*, verbessert sich das Diskriminationsvermögen der Probanden, je mehr der Parameter, die aus den natürlichen Gipfelgestalten von Frage-Äußerungen extrahiert wurden, in die getestete Kontur integriert sind. Für die hier untersuchte intonatorische Differenzierung zwischen Frage und Aussage muß dabei der Parameter der Anstiegsneigung als der stärkere *cue* im Vergleich zur Gipfelhöhe gewichtet werden: In der Identifikation zeigt sich der Einfluß der Neigungsvariable auf das Urteil der Probanden als hoch signifikant (*,000*) und in der Diskrimination als ein signifikanter Trend (*,039*). Dieses Ergebnis unterstützt die eingangs gestellte **Hypothese 3.** und widerlegt die **Hypothese 2.** (s. Unterkapitel 5.1.).

Des weiteren werden in Abhängigkeit von der Gipfelgestalt unterschiedliche Stimuli schärfer diskriminiert. Die Diskriminationsgrenze für die Gipfelgestalten *hoch/steil* und *tief/steil* liegt bereits zwischen der frühen und der mittleren Gipfelpositionierung (-30 vs. +30 ms bzw. 0 vs. +60 ms), während für die Kontur *tief/flach* eine Grenze zwischen mittel und spät (+30 vs. +90 ms) diskriminiert wird. Die Maxima in den Diskriminationskurven entsprechen jedoch in allen Fällen den Anstiegen der *Frage*-Urteile auf ca. 50% in der Identifikation des jeweiligen zweiten Stimulus im Paar: Die Diskriminationsschärfe fällt also in den Kategorienübergang (vgl. Repp, 1984). Für die Gipfelkontur *tief/flach* sind die relevanten

Änderungen im Identifikations- wie Diskriminationsverhalten der Probanden allerdings nur sehr schwach ausgeprägt.

Die Ergebnisse prüfstatischer Analysen zeigen in erster Linie, daß für die Gesamtheit der getesteten Konturen eine hoch signifikante Änderung (*,000*) in der Frage-Identifikation bei der Gipfelverschiebung von der Mitte in das letzte Drittel des Akzentvokals (+60 zu +90 ms) vorliegt. Der Identifikationssprung wird von einigen signifikanten Befunden in der Diskrimination begleitet: Hier findet über alle Gipfelgestalten ein Rückgang der *verschieden*-Urteile (*,010*) vom Stimuluspaar +30/ +90 zu den Stimuli +60 vs. +120 statt. Eine weitere signifikante Abnahme der *verschieden*-Urteile (*,003*) erfolgt vom Stimuluspaar +60/+120 zum Paar +90/+150. Dies läßt darauf schließen, daß bei der mittleren Gipfelpositionierung von +60 ms eine Kategoriengrenze vorliegt und die spät positionierten Gipfel (+90 bis +150 ms) aller getesteten Gestalten bei der Diskrimination der gleichen Kategorie zugeordnet werden.

Die Diskrimination der steil vs. flach ansteigenden Gipfelkonturen (Faktor Neigung) weist zwei signifikante Befunde (*,003* und *,001*) auf, die auf gegenläufigen Effekten bei der Bewertung der Stimuluspaare -30 vs. +30 ms und +30 vs. +90 ms beruhen: Während die Diskrimination der steil ansteigenden Gipfelkonturen beim mittleren Gipfel (+30 ms) stark ausfällt und von einer Abschwächung beim späten Gipfel (+90 ms) gefolgt wird, ist sie bei Gipfelkonturen mit dem flachen Anstieg beim Stimuluspaar -30/+30 schwach ausgeprägt, steigt jedoch zum späten Gipfel an. Diese Ergebnisse werden allerdings nicht durch eine Entsprechung in der Inferenzstatistik zur Identifikation bekräftigt.

Aufgrund der oben angeführten Überlegungen wird die **Hypothese 1.** als bestätigt angenommen. Die späte Gipfelposition (ab der zweiten Hälfte des Akzentvokals) ist ein dominanter *cue* in der intonatorischen Kodierung der Frage-Kategorie, jedoch kann hierbei eine Wechselwirkung zwischen Position und Neigung auftreten: Gipfel mit einem flach verlaufenden Anstieg müssen im Vergleich zu steil ansteigenden Gipfelkonturen später in Relation zum Akzentvokal positioniert werden, um ein Frage-Perzept zu erzeugen. Zu einem vergleichbaren Ergebnis führte die psychophonetische Studie von Niebuhr (2003), der sich mit der akustischen Ausprägung der Kategorien des frühen vs. mittleren Gipfels im Deutschen beschäftigte (zu deren Semantik s. Abschnitt 3.2.3). Dort konnte festgestellt werden, daß eine Gipfelgestalt mit steilem Abstieg eine signifikante Auswirkung auf die Perzeption eines mittleren Gipfels ausübt. Eine Gipfelgestalt mit steilen An- und Abstiegen ruft bereits bei einer frühen Gipfelpositionierung vor dem Vokalonset den perzeptiven Eindruck eines mittleren Gipfels hervor. Auch im Ungarischen wurde eine Interaktion zwischen Gipfelgestalt

und Timing bei der Perzeption der Kategorien Frage vs. Aussage beobachtet (vgl. Gósy und Terken, 1994; Unterkapitel 4.2.1.): Bei einem mittleren Gipfel kann eine kontinuierliche Neigung des Anstiegs die Frage-Wahrnehmung abschwächen, der kontinuierliche Gipfelabstieg unterstützt dagegen die Perzeption der Äußerung als Frage. Solche Ergebnisse weisen auf eine perzeptive Sensibilität gegenüber den Tonbewegungen in Relation zum Akzentvokal bzw. den tontragenden Segmenten der Akzentsilbe hin, wobei sowohl das Timing als auch die Gestalt der steigend-fallenden Gipfelkonturen relevant erscheinen.

Die oben angeführten ähnlichen Ergebnisse aus drei unterschiedlichen Sprachen sind als ein Hinweis auf eine psychophonetische Erscheinung aufzufassen, die sprachabhängig auf verschiedene Art funktionalisiert werden kann (vgl. auch Kohler, 1991a: 156; Unterkapitel 3.2.3.): Intonationsverläufe über Akzentsilben werden in Verbindung mit Sprachkonzepten gebracht, wobei sich prototypische Konturen etablieren (vgl. unten) und sprachspezifische perzeptive *cues* bilden. So fungieren neben einem nicht-frühen Gipfeltiming im Russischen ein steiler Anstieg und im Ungarischen eine relativ hohe Gipfellage (Gósy und Terken, 1994) als dominante *cues* für ein Frage-Perzept. Aufgrund der Daten zum neapolitanischen Italienisch (vgl. D'Imperio und House, 1997; Abschnitt 4.2.2.) kann die Hypothese aufgestellt werden, daß hier eher ein kontinuierlicher Anstieg bei einem nicht-frühen Gipfel die Frage-Wahrnehmung fördert. Im Fall von abweichenden, unprototypischen Realisierungen können kompensatorisch andere *cues* (wie z.B. Gipfelhöhe im Russischen oder Neigung der An- und Abstiegsbewegungen im Ungarischen) in der Perzeption stärker gewichtet werden. Die dargelegte Auffassung steht im Einklang mit der Hypothese einer hochgradigen Flexibilität des menschlichen Perzeptionssystems, welches über eine Fähigkeit verfügt, sich den situativen Erfordernissen anzupassen und somit eine erfolgreiche Kommunikation zu ermöglichen (s. Reinholt Petersen, 1986; Abschnitt 2.1.3.).

In der linguistischen Forschung wurde häufig darauf hingewiesen, daß sprachliche Kategorien prototypisch strukturiert sind, wobei ein Prototyp ein zentrales bzw. repräsentatives Mitglied einer Kategorie darstellt (vgl. Saeed, 1997: 37). In dem durchgeführten Experiment konnte ein Prototypikalitätseffekt beobachtet werden: Das stilisierte Original der Frageäußerung wurde in allen Tests zu 97,2% bis 100% der Fälle als Frage bewertet. Dieser Effekt trat ebenso bei der Beurteilung der manipulierten Stimuli der vierten Experimentalreihe auf (93-95% der *Frage*-Urteile): Hierbei handelte es sich um spät positionierte Gipfel mit der Gestalt *hoch/steil*. Dieses Ergebnis kann im Sinne der eingangs gestellten **Hypothese 4**. (s. Unterkapitel 5.1.) so interpretiert werden, daß die Unterscheidung zwischen Aussage und Frage im Russischen bei äußerungsfinalen Fragefokus ausschließlich von drei Para-

metern (Höhe, Aufstiegsneigung und zeitliche Positionierung des F0-Gipfels) getragen wird. Denn mit diesen drei Parametern kann bereits ein Hörerurteil erreicht werden, das der Beurteilung natürlichsprachlicher Fragekontur beinahe vollständig entspricht. Zur prototypischen Intonationskontur der Aussage ist anzumerken, daß sie durch eine frühe Gipfelposition vor dem Onset des Akzentvokals und eine flache Anstiegsneigung charakterisiert wird. Eine Kontur mit steil verlaufendem Gipfelanstieg ist dagegen untypisch und ruft bereits bei einer frühen Gipfelpositionierung ein weniger eindeutiges *Aussage*-Urteil hervor.

Ein zusätzliches ausgeprägtes Diskriminationsmaximum, welches speziell bei den Stimuli –60 vs. 0 der Gipfelgestalten *hoch/flach* und *tief/steil* registriert wurde (hier liegt eine signifikante Wechselwirkung ( $,004$ ) unter den Gruppenfaktorbedingungen auf dem Meßwiederholungsfaktor vor), deutet möglicherweise auf die Präsenz einer zusätzlichen Kategorie im hergestellten Kontinuum hin. Eine solche Interpretation wird durch die Ergebnisse der durchgeführten Vortests gestützt (s. Abschnitt 5.3.5.): Die intuitiven Reaktionen der Probanden zeigen, daß ein hoher bzw. steil ansteigender früher Gipfel als ein Ausdruck einer kontrastierenden Hervorhebung verstanden werden kann. Auch in anderen Sprachen wurde Kontrast bzw. Emphase in Beziehung zu Extension und Höhe von Gipfelkonturen gesetzt (vgl. Gussenhoven (2002) für einen Überblick). Die Klärung dieser Fragestellung für das Russische bedarf jedoch einer gesonderten Untersuchung. In diesem Zusammenhang wäre die Auswertung der Antworten derjenigen Versuchspersonen interessant, die aufgrund ihrer abweichenden Bewertung des Frage-Originals als *Kontrast* aus dem Datenkorpus der vorliegenden Arbeit ausgenommen wurden (vgl. Abschnitt 5.3.1.).

Mit der vorliegenden Arbeit erfolgte eine erste Annäherung an eine „*study of interpretations and meaning of the established types of pitch accent*“, was Odé in ihrer Perzeptionsanalyse des Russischen der zukünftigen Forschung vorbehielt (vgl. Odé, 1989: 39). Die Ergebnisse der durchgeführten Perzeptionsexperimente legen dabei nahe, daß die von Odé (1989) aufgestellte Phonologie der perzeptiv relevanten Intonationskategorien erweitert werden muß: Der hier festgestellte, für die Frage-Kategorie prototypische Konturtyp mit einem steilen Anstieg, späten Timing und großen Intervall (impliziert durch die Gipfelhöhe) (*Rl+* in Termini von Odé) findet sich nicht unter den beschriebenen *types of pitch accents* (vgl. Unterkapitel 4.1.). Weiterhin ist anzunehmen, daß in der Intonationsphonologie des Russischen der Parameter der Anstiegsneigung in Gipfelkonturen als distinktiv berücksichtigt werden muß. Dies spricht gegen die Auffassung von Odé (1989), daß die Neigung in steigenden *pitch accents* nicht als ein unabhängiges Merkmal auftritt. Der Parameter Neigung wird in der Odé’schen Intonationsphonologie von Intervall und Timing der perzeptiv

relevanten steigenden F0-Bewegung dominiert. Das Intervall der F0-Bewegung, welches bei Odé den Parameter Höhe inkorporiert, erweist sich im Zusammenhang mit dem hier untersuchten kategorialen Unterschied jedoch als ein untergeordnetes Merkmal. Die Parameter des Timings und der Neigung können zwar innerhalb eines gewissen Rahmens (s. oben) interagieren, scheinen jedoch bei der Unterscheidung zwischen den Kategorien der Frage, der Aussage und des Kontrasts distinktiv zu sein. Diese Annahme sollte in weiterführenden Untersuchungen verfolgt werden.

Die vorliegende Untersuchung ist im Kontext der modernen Intonationsforschung zum Russischen ein Novum: Der hier präsentierte Ansatz liefert ein Beispiel für eine Intonationsphonologie, die auf einer Analyse von Produktionsdaten basiert und eine perzeptorische Überprüfung der aufgestellten Kategorien unter Berücksichtigung der semantischen Interpretationen durch den Hörer anstrebt. Ein solches methodisches Vorgehen wurde bereits erfolgreich bei der Modellierung der deutschen Intonation eingesetzt (vgl. Kohler, 1991b,c). Die vorliegende Arbeit beschränkte sich zwar auf einen klein umrissenen Bereich ausgewählter sprachlicher Kategorien, konnte jedoch zeigen, daß die Odé'sche Zusammenstellung der distinktiven intonatorischen Merkmale des Russischen erweitert werden muß. Dies weist deutlich darauf hin, daß bei der Aufstellung einer Intonationsphonologie die semantisch-pragmatische Komponente nicht fehlen darf.

## **6.2. Universale Tendenzen im Sprachvergleich**

Die Ergebnisse der durchgeführten Untersuchung scheinen mit der sprachuniversalen Tendenz, Entscheidungsfragen durch final steigende Konturen zu markieren (vgl. *The Strong Universalist Hypothesis* von Ladd, 1981), unvereinbar zu sein. Die Universalitätshypothese in dieser Form wurde von Ladd (1981) zurecht kritisiert. Jedoch kann ein Bestreben, im Bereich der so grundlegenden Sprachkategorien wie Frage und Aussage universale Tendenzen aufzuspüren, nicht als ein „Prokrustes-Unterfangen“ abgetan werden (vgl. Ladd, 1981: 389).

Im Hinblick auf die Ergebnisse der einzelsprachlichen bzw. komparativen Untersuchungen zur Produktion und Perzeption der intonatorischen Muster in Fragen und Aussagen (s. unten) wird deutlich, daß die Universalitätshypothese anders formuliert werden sollte: Sie ist eher als eine Tendenz aufzufassen, höhere bzw. steigende Konturen zum Ausdruck der Frage und im Kontrast dazu tiefere bzw. fallende Konturen zum Ausdruck der Aussage zu verwenden, und zwar ohne einen festen Bezugspunkt in der Äußerung, denn dies

verstieße gegen das konventionelle Prinzip des sprachlichen Gebrauchs (vgl. de Saussure, 1931).

Bereits Hermann (1942) stellte in seiner Studie über 165 Sprachen der Welt fest, daß die Entscheidungsfrage ohne Ausnahme durch einen hohen Ton gekennzeichnet ist. Dabei beschränke sich das Vorkommen dieses hohen Tons im Sprachvergleich nicht nur auf das Satzende, sondern könne auch an anderen Stellen der Äußerung auftreten: „*Überhaupt ist nicht nur am Anfang und Schluß, sondern im ganzen Satz die Tonlage der Frage sehr häufig höher als die der Aussage*“ (Hermann, 1942: 331). In der Perzeptionsforschung wurde häufig vermerkt, daß in steigend-fallenden Gipfelkonturen höhere, steigende Muster über der Akzentsilbe als fragend wahrgenommen werden. Dies zeigt sich nicht nur in den Sprachen wie dem Russischen, dem Ungarischen und dem neapolitanischen Italienisch, welche die Unterscheidung zwischen höheren, nicht-frühen und tieferen, frühen Gipfeln als einen Teil des Sprachkodes verwenden (vgl. Unterkapitel 4.2.). Diese Tendenz konnte auch bei anderen Hörern mit unterschiedlichem Sprachhintergrund festgestellt werden: So fanden Gussenhoven und Chen (2000) heraus, daß chinesische, holländische ebenso wie ungarische Hörer dazu neigen, höhere und späte Gipfel (in trisyllabischen Nonsens-Äußerungen) als Fragen zu interpretieren. Im Deutschen kann eine Längung der ersten Hälfte des Akzentvokals, über welche ein Anstieg zu einem mittleren Gipfel stattfindet, die Wahrnehmung eines fragenden Klangs erzeugen, wobei die Verlängerung des Anstiegs als eine perzeptive Hervorhebung des höheren Tons über dem akzentuierten Vokal interpretiert werden kann (Dombrowski, 2003).

Den universellen Charakter eines derartigen Zusammenhangs zwischen Form (ein relativ hoher Ton) und Funktion (Fragesignal) erkannte schon Hermann (1942). Er suchte eine Erklärung für die Entstehung des hohen Fragetons in der „*Psychologie des menschlichen Innenlebens*“, wobei ein Hochtton als besser vernehmlich, die Deutlichkeit des Gesagten erhöhend und Aufmerksamkeit erweckend gedeutet wurde (vgl. Hermann, 1942: 343ff.). Ein anderer Ansatz versucht die sprachübergreifende Tendenz zur Verwendung von höheren Tönen in Fragen und tieferen Tönen in Aussagen aus ethologischer Perspektive zu klären (*frequency code*, Ohala, 1984). Eine zentrale Annahme der Theorie des *frequency code* betrifft die symbolische Natur der tonalen Höhe vs. Tiefe, welche primär auf die Andeutung einer Stärkerelation zurückgeführt werden kann: Für den Ausdruck der Dominanzverhältnisse wird bei Menschen sowie anderen Spezies unter anderem der Zusammenhang zwischen der Körpergröße und der Larynxgröße ausgenutzt, wobei bekanntlich die Größe des Larynx und die Beschaffenheit der Stimmlippen im Wesentlichen die Höhe des produzierbaren Stimmtons beeinflussen, so daß kleine Laryngen (mit kürzeren, dünneren Stimmlippen) generell höhere

Töne erzeugen (vgl. Unterkapitel 2.1.2.). Ein tiefer bzw. fallender Ton kann somit sekundär als ein Ausdruck von Stärke, Sicherheit, Dominanz sowie Aggression interpretiert werden, während einem hohen bzw. steigenden die Bedeutung von Schwäche, Unsicherheit, Unterlegenheit oder auch Freundlichkeit zukommt. Im Sinne des *frequency code* wird durch die Markierung einer Frage mit einem hohen Ton die Unsicherheit des Sprechers bezüglich des Geäußerten signalisiert. Der Sprecher bekundet seine Unterlegenheit gegenüber dem Hörer, der über die benötigte Information verfügt und sie in Form einer Antwort geben kann (s. Abschnitt 2.2.1. zur Semantik und Pragmatik von Fragen). Aus dieser Sicht kann also die Universalität der Verwendung eines hohen Tons auf seine außersprachlichen affektiven Bedeutungen zurückgeführt werden.

Die konventionelle Natur sprachlicher Ausdrucksmittel kommt im Bereich der intonatorischen Markierung der diskutierten Handlungstypen dadurch zum Vorschein, daß der Kontrast zwischen höheren Tönen in Fragen und tieferen in Aussagen im Sprachvergleich auf verschiedene Art ausgedrückt wird: Dies kann (1) auf der Ebene der Äußerung wie z.B. im Chinesischen (Yuan et al., 2002) und Dänischen (Thorsen, 1979, 1985), (2) auf der Akzentsilbe in steigend-fallenden Gipfelkonturen (im Russischen, Ungarischen) sowie in final steigenden Konturen (z.B. im australischen Englisch; vgl. Fletcher und Harrington, 2001), (3) am Äußerungsende wie z.B. im Hindi, Maya und Tschechischen (nach Bolinger, 1978: 502) stattfinden. Des weiteren können die Intonationsmuster der Entscheidungsfragen im zwischensprachlichen Vergleich unterschiedlich spezifiziert sein: (1) So wird in Fragen die langsam absteigende obere Deklinationslinie angehoben (Chinesisch) oder gar aufgehoben (Dänisch); (2) In steigend-fallenden Konturen mit einem Gipfelanstieg über der Akzentsilbe können z.B. Gipfelhöhe (Ungarisch) oder Anstieg des Gipfels (Russisch) stärker gewichtet werden; (3) Bei äußerungsfinalen Talkonturen kann der Höhe des Anstiegs unterschiedliche Relevanz zukommen, wobei sowohl tief steigende als auch hoch steigende Konturen zu beobachten sind (wie im Chrau und Tagalog vs. im Portugiesischen und Aserbaidshaischen; nach Cruttenden, 1997: 157).

Die Konventionalität des Sprachgebrauchs und die Besonderheiten der Sprachentwicklungsgeschichte treten auch bei denjenigen Sprachbeispielen in Erscheinung, welche mit der festgestellten universalen Tendenz unvereinbar sind: So kommt z.B. in den süd-nigerianischen Tonsprachen Degema, Jsoko und Jzon ein tiefer Ton als Marker für Entscheidungsfragen vor (Williamson, 1979). Die Verwendung dieses Tons scheint auf eine tief-tonige Fragepartikel zurückzuführen, die im Laufe des Sprachwandels abgelegt, deren Ton jedoch behalten wurde.

Es gibt jedoch weitere Hinweise für die Annahme, daß der *frequency code* zum menschlichen „*genetic make-up*“ (Ohala, 1984: 10) gehört und somit eine universale Quelle affektiver Bedeutungen und der daraus resultierenden Verwendungen eines hohen Tons darstellt. Im Französischen können z.B. die Konturen in Aussagen mit einem finalen Anstieg enden, wenn der Sprecher unsicher ist (Cruttenden, 1981). Im amerikanischen Englisch werden tief verlaufende oder auch final stark fallende Intonationskonturen als Ausdruck der Dominanz und Selbstsicherheit interpretiert (Ohala, 1984). Im Japanischen wirken final steigende Konturen als ein Zeichen für Sympathie und Teilnahme (Abe, 1955). Hier werden weiterhin Gipfelkonturen mit einer verzögerten Realisierung des F0-Maximums in Relation zur Akzentsilbe (*F0-peak delay*, auch *ososagari*) bei einer femininen Sprechweise beobachtet, welche sich durch eine stärkere Einbeziehung des Hörers ins Gespräch auszeichnet (Hasegawa und Hata, 1995). An dieser Stelle sei auch das Phänomen der sogenannten Ammensprache (*motherese*) erwähnt. Dieser spezielle Sprechstil, der für die kommunikative Interaktion von Eltern mit Säuglingen charakteristisch ist, wurde in sehr unterschiedlichen Sprachgemeinschaften (z.B. im Deutschen, Englischen ebenso wie im Mandarin-Chinesisch) beobachtet (Papoušek et al., 1991). Die Sprechweise der Erwachsenen zeichnet sich hierbei durch eine häufige Verwendung hochfrequenter und steigender Intonationskonturen zur Dialoganregung aus. Niederfrequente fallende Sprechmelodien werden dagegen zum Beruhigen eingesetzt.

Die oben angeführten Beispiele sprechen für die Annahme der Theorie des *frequency code*. Die Hypothese von Hermann (1942), ein hoher Ton sei deutlicher und besser vernehmlich, läßt sich dagegen aus psychophonetischer Sicht nicht belegen (vgl. Unterkapitel 2.1.3.).

Wie der Sprachvergleich zeigt, sollte nicht allein der äußerungsfinale Unterschied zwischen steigenden und fallenden Konturen als ein sprachuniversales intonatorisches Merkmal zur Differenzierung der Sprecherabsichten betrachtet werden (Ladd, 1981), sondern vielmehr der Gegensatz *relative tonale Höhe* vs. *Tiefe* (vgl. auch Nikolajeva, 1989: 5). Die Markierung von Fragen vs. Aussagen durch einen äußerungsfinalen Kontrast (steigend vs. fallend), der sich für die Satzmoduszeichnung in den meisten Sprachen der Welt etablierte (Bolinger, 1978), kann möglicherweise dadurch erklärt werden, daß finale Tonänderungen für Hörer leichter wahrzunehmen und zu verarbeiten sind. So zeigt eine psychoakustische Untersuchung von 't Hart (1981), daß einige Hörer (mit holländischem Sprachhintergrund) dazu neigen, eher die Abweichungen in den finalen Tonlagen zweier Äußerungskonturen zu differenzieren, als sich auf die unterschiedlichen Intervalle in den internen melodischen

Äußerungsstrukturen zu konzentrieren. Dies mag mit den begrenzten Möglichkeiten der menschlichen Gedächtniskapazität zusammenhängen: Wie eine weitere psycholinguistische Studie ergab, wird die Tonhöhe eines Sprachsignals im sensorischen Gedächtnis lediglich für ca. 2 s gespeichert (Helfrich, 1985). Dieses Ergebnis gibt einen Hinweis darauf, daß ein Kontrast zwischen hoch und tief in finaler Position einfacher zu diskriminieren ist als in medialer Position von längeren Äußerungen. Für die durchgeführte Untersuchung ist dieser Hinweis insofern relevant, als die vorliegenden Befunde nicht dahingehend verallgemeinert werden können, daß sie die alleinigen intonatorischen Ausdrucksmittel für die Satzmodusdifferenzierung im Russischen darstellen. Aus diesem Grund erscheint eine weiterführende Untersuchung mit der Fragestellung sinnvoll, ob die hier festgestellte formale Unterscheidung auch bei einem äßerungsinitialen bzw. –medialen Fragefokus die relevanten Handlungstypen Frage vs. Aussage im Russischen ausreichend kodiert oder ob eventuell in diesen Fällen zusätzliche finale *cues* notwendig werden. Bei einer nicht-finalen Platzierung des Fragefokus wird möglicherweise auch der in dieser Untersuchung nicht kontrollierte Parameter der Abstiegsneigung relevant.

### **6.3. Kritische Bemerkungen zum durchgeführten Experiment**

In dem durchgeführten Experiment wurde ein Konturparameter, nämlich die Lage des Endpunktes der gesamten Äußerungskontur, nicht berücksichtigt. Die Äußerungskontur der Teststimuli mit den späten Gipfelpositionierungen (+120 und +150) endete insgesamt geringfügig höher, als die der Stimuli mit einem früher im Kontinuum positionierten Gipfel (vgl. Abschnitt 5.2.2.4.). Dies ist eventuell ein zusätzlicher Parameter, welcher die Wahrnehmung einer Frage bei den Stimuli mit einer späten Gipfelposition im Experiment unterstützen kann. In den bisherigen Perzeptionsuntersuchungen zum Russischen wurde ebenfalls angenommen, daß die Endlage der fallenden postakzentualen F<sub>0</sub>-Bewegung für die Wahrnehmung der steigenden Akzente relevant ist (Odé, 1989; s. auch 4.1.).

Problematisch erscheint auch die in dieser Untersuchung getroffene Auswahl an Versuchspersonen, welche stark unterschiedliche Herkunftshintergründe aufweisen (s. Unterkapitel 5.2.4.). Durch die potentiellen, unberücksichtigt gebliebenen Dialektunterschiede ist möglicherweise die starke Streuung zu erklären, die bei der prüfstatistischen Analyse beobachtet wurde. Für die weitere Forschung wäre daher wünschenswert, Datenerhebungen innerhalb von homogenen Varietäten des Russischen durchzuführen. Angesichts einer relativ

geringen Anzahl der in den Tests beteiligten Personen müssen die vorliegenden Ergebnisse der Arbeit als eine Tendenz gesehen werden, deren Generalisierung für die Population der russischen Sprecher jedoch bis auf weitere Datenerhebungen eingeschränkt bleibt.

#### **6.4. Vorschläge für weiterführende Forschung**

Die Ergebnisse der durchgeführten Vortests (s. Abschnitt 5.3.5.) legen nahe, daß der Ausdruck einer kontrastierenden Hervorhebung im Russischen durch ähnliche intonatorische Merkmale wie die Entscheidungsfrage gekennzeichnet wird. Die intonatorische Markierung dieses Gegensatzes wurde bisher nur unzureichend erforscht (vgl. auch 3.1.). Anzunehmen ist, daß bei einem Kontrastakzent im Russischen in erster Linie die Positionierung des Gipfels und die Gipfelhöhe, gefolgt von der Neigung des Gipfelanstiegs als wichtige *acoustic cues* fungieren. Auf dem zeitlichen Kontinuum der Synchronisierung des Gipfels mit der Akzentsilbe kann der Gegensatz zwischen früheren und späteren Gipfelpositionen Kontrastakzent vs. Entscheidungsfrage differenzieren, wobei hier anscheinend eine Quelle für dialektale Variation liegt (vgl. 5.3.5.). Wie oben (Abschnitt 6.3.) erwähnt, sind in der Zukunft dialektinterne Datenerhebungen notwendig.

Eine weitere Fragestellung betrifft die Position des Fragefokus innerhalb einer Äußerung. Im durchgeführten Experiment wurden zwei mögliche Realisierungen von Äußerungen mit Frageintention nicht thematisiert: (1) Das in der Frage fokussierte Wort kann nicht nur final, sondern auch medial oder am Anfang einer Äußerung vorkommen, wobei möglicherweise zusätzliche *acoustic cues* zur intonatorischen Markierung der Frage erforderlich werden (vgl. 6.2.); (2) Natürlichsprachliche Äußerungen weisen häufig mehr als nur einen Akzent auf, was auch im Fall von Fragen möglich ist. So wurde für das Ungarische gezeigt, daß die Präsenz einer zusätzlichen Akzentkontur in der Äußerung ihre Wahrnehmung als Frage negativ beeinflusst (Gósy und Terken, 1994; Unterkapitel 4.2.1.). In der einschlägigen Literatur zum Russischen finden sich leider kaum Angaben über die Intonationsstruktur von komplexen Frage-Äußerungen (vgl. 3.1.).

In der alltäglichen Kommunikation wirken sich mehrere Faktoren auf das Sprachverstehen aus, wobei unter anderem der Kontext eine große Rolle spielt. In dieser Hinsicht schuf das durchgeführte Experiment eine idealisierte Sprachsituation, in welcher die Hörer eine isoliert präsentierte Äußerung zu beurteilen hatten. Um realitätsnahe Experimentbedingungen zu schaffen, soll das Paradigma der kategorialen Wahrnehmung durch weitere

Forschungsparadigmen erweitert werden. Dies kann beispielsweise durch die Aufnahme von kurzen Dialogsequenzen ins Perzeptionsexperiment erfolgen. Unter der Berücksichtigung dessen, daß die sprachlichen Kategorien prototypisch strukturiert sind, würde solch eine Vorgehensweise unter anderem eine formale Trennung zwischen kontextuell tolerierbaren und prototypischen Frage-Ausdrücken ermöglichen.

## 7. Zusammenfassung und Ausblick

Die vorliegende empirische Perzeptionsuntersuchung zum Intonationssystem des Russischen beschäftigte sich mit der Fragestellung, welche akustischen Schlüsselreize (*acoustic cues*) im Bereich der Intonation für russische Hörer relevant sind, um die Intention einer Äußerung als eine Entscheidungsfrage oder als eine Aussage zu interpretieren. Es konnte gezeigt werden, daß sich der intonatorische Kontrast zwischen den Kategorien der Frage und Aussage in äußerungsfinaler Position auf die Parameter Gipfelhöhe, Anstiegsneigung und Gipfelpositionierung beläuft. Als besonders relevante *acoustic cues* stellten sich die Positionierung des Gipfels und die Neigung des Anstiegs heraus, wobei eine späte Gipfelposition und eine steile Neigung starke *perceptive cues* für die Wahrnehmung der Frage bilden. Die *perceptive Relevanz* der Gipfelhöhe für die Kodierung des hier untersuchten semantisch-pragmatischen Aspektes ist dagegen geringer einzuschätzen. Prototypischerweise wird die Entscheidungsfrage bei finalem Fragefokus durch ein spätes Timing, relativ hohe Gipfellage und eine steile Neigung des Gipfelanstiegs kodiert. Die prototypische Intonationskontur der Aussage wird dagegen durch ein frühes Timing, ein geringeres Gipfelmaximum und eine flache Anstiegsneigung gekennzeichnet.

Die Ergebnisse dieser Untersuchung unterstützen eine sprachübergreifende Tendenz, höhere bzw. steigende tonale Konturen zur Markierung von Fragen einzusetzen, die im Rahmen der Theorie des *frequency code* als Ausdruck eines universellen, außersprachlichen Wissens erklärt werden kann (Ohala, 1984).

Mit der vorliegenden Arbeit wurde der erste Versuch in Richtung einer auf Perzeption basierenden Intonationsphonologie des Russischen unternommen, die einen Bezug zu semantisch-pragmatischen Kategorien aufbaut. Dieser methodische Ansatz stellte sich im Vergleich zu der Perzeptionsforschung, welche die Interpretation der Äußerungen durch den Hörer eliminiert, als perspektivenreicher heraus. Die Kombination aus den sich ergänzenden Forschungsbereichen der Produktion und Perzeption unter Einbeziehung der semantisch-pragmatischen Aspekte ermöglicht es, prototypische Intonationskonturen sowie die Relevanz ihrer einzelnen Merkmale zu bestimmen und die Perzeptionsmechanismen bei Ambiguität der produzierten Sprachsignale aufzudecken.

Außer in der phonologischen Theoriebildung sind die Ergebnisse der durchgeführten Untersuchung für die praktischen Belange des Fremdsprachenunterrichts anwendbar. Die Intonationskonturen in russischen Entscheidungsfragen wurden häufig als äußerst schwierig

bezeichnet, insbesondere für Lernende mit Muttersprachen, in denen für diese Art von Frage eine final steigende Talkontur (wie z.B. im Deutschen) charakteristisch ist (vgl. Wenk, 1975: 188). Die Arbeitsergebnisse können hierbei zur Klärung der Variation unter den zulässigen Realisierungen der Intonationskonturen in Frage-Äußerungen verhelfen, um einen erfolgreichen kommunikativen Austausch in der Fremdsprache Russisch vollziehen zu können.

Das durchgeführte Experiment zeigt mehrere weiterführende Fragestellungen auf, welche die Abgrenzung zwischen den intonatorischen Ausdrücken von Frage vs. Kontrast und die Intonationsmerkmale der Entscheidungsfragen in komplexeren Äußerungen umfassen. Durch die Ergebnisse des Experiments wird deutlich, daß bei den zukünftigen Untersuchungen dialektale Unterschiede im Bereich der russischen Intonation beachtet werden müssen. Außerdem sollten die Methoden der perzeptorischen Intonationsforschung unter Berücksichtigung der realen Kommunikationsbedingungen (wie z.B. kontextuelle Einbettung der Aussage-Frage-Sequenzen bei einem Dialog) erweitert werden.

## Literaturliste

- Abe, I. (1955): Intonational pattern of English and Japanese. In: *Word* 2. S. 386-398.
- Aedrians, L. M. H. (1991): Ein Modell deutscher Intonation. Eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzveränderungen im gelesenen Text. Ph. D. Diss. Eindhoven: Technische Universität.
- Altmann, H. (1987): Zur Problematik der Konstitution von Satzmodi als Formtypen. In: Meibauer, J. (ed.) *Satzmodus zwischen Grammatik und Pragmatik*. Tübingen: Max Niemeyer. S. 22-56.
- Berg, J. van den (1968): Mechanism of the larynx and the laryngeal vibrations. In: B. Malmberg (ed.) *Manual of Phonetics*. Amsterdam: North Holland. S. 278-308.
- Bondarko, L. V. (1998): Fonetika sovremennogo russkogo jazyka. [Phonetik der modernen russischen Sprache.] St.-Petersburg.
- Bolinger, D. L. (1972): Accent is predictable (if you are a mind reader). In: *Language* 48. S. 633-644.
- Bolinger, D. L. (1978): Intonation across languages. In: Greenberg et al. (eds.) *Universals of human language. Vol. 2 (Phonology)*. S. 471-525.
- Bortz, J. (1979): Lehrbuch der Statistik für Sozialwissenschaftler. Berlin: Springer-Verlag.
- Brosius, F. (2002): SPSS 11. Bonn: mit Verlag.
- Brosnahan, L. F., Malmberg, B. (1970): Introduction to phonetics. Cambridge: Heffer & Sons.
- Bryzgunova, E. A. (1977): Zvuki i intonatsija russkoj retshi. [Laute und Intonation der russischen Sprache]. Moskva: Russkij jazyk.
- Bryzgunova, E. A. (1980): Intonatsija. [Intonation.] In: Shviedova, N. J. (ed.) *Russkaja grammatika, Bd. 1*. Moskva: Nauka. S. 96-122.
- Bryzgunova, E. A. (1984): Emotsionaljno-stilistitsheskije razlitshija russkoj zvu-tshashnej retshi. [Emotionale und stilistische Unterschiede in der russischen Spontansprache.] Moskva.
- Bühler, K. (1934): Sprachtheorie. Stuttgart.
- Cruttenden, A. (1981): Falls and rises: meanings and universals. In: *Journal of Linguistics* 17. S. 77-91.
- Cruttenden, A. (1997): Intonation. Cambridge: Cambridge University Press.

Crystal, D. (1969): *Prosodic Systems and Intonations in English*. Cambridge University Press: Cambridge.

Dombrowski, E. (2003): Semantic features of accent contours: Effects of F0 peak position and F0 time shape. In: *Proceedings of the XV International Congress of Phonetic Sciences*. Barcelona. (im Druck)

Essen, O. von (1964): *Grundzüge der hochdeutschen Satzintonation*. Ratingen: Henn.

Féry, C. (1993): *German Intonation Patterns*. Tübingen: Max Niemeyer.

Fletcher, J., Harrington, J. (2001): High rising terminals and fall-rises in Australian English. In: *Phonetica* 58(4). S. 215-229.

Fries, C. C. (1964): On the intonation of ‚Yes-No‘ questions in English. In: Abercrombie, D. et al. (eds.) *In honours of Daniel Jones*. London: Longmans. S. 242-254.

Gårding, E. (1983): A generative model of intonation. In: Cutler, A., Ladd, D. R. (eds.) *Prosody: models and measurements*. Berlin: Springer Verlag. S. 11-25.

Gartenberg, R., Panzlaff-Reuter, C. (1991): Production and Perception of F0 Peak Patterns in German. In: K. J. Kohler (ed.) *Studies in German Intonation. Arbeitsberichte des Instituts für Phonetik Universität Kiel (AIPUK 25)*. S. 31-113.

Gósy, M., Terken, J. (1994): Question marking in Hungarian: timing and height of pitch peaks. In: *Journal of Phonetics* 22. S. 269-281.

Gussenhoven, C. (2002): Intonation and Interpretation: Phonetics and Phonology. In: *Proceedings Speech Prosody 2002*. Aix-en-Provence. URL3.

Gussenhoven, C., Chen, A. (2000): Universals and language-specific effects in the perception of question intonation. In *Proceedings of ICLSP 6 (2)*, Beijing. S. 91-94.

Hadding-Koch, K., Studdert-Kennedy, M. (1964): An experimental study of some intonation contours. In: *Phonetica* 11, S. 175-185.

Halliday, M. A. K. (1967): *Intonation and Grammar in British English*. Paris: Mouton.

Handel, S. (1989): *Listening. An Introduction to the Perception of Auditory Events*. Cambridge: The MIT Press.

’t Hart, J. (1981): Differential sensitivity to pitch distance, particularly in speech. In: *Journal of Acoustical Society of America* 69 (3). S. 811-821.

’t Hart, J., Collier, R., Cohen, A. (1990): *A perceptual study of intonation. An experimental-phonetic approach of speech melody*. Cambridge: Cambridge University Press.

Hasegawa, Y., Hata, K. (1995): The function of F0-peak delay in Japanese. In: *Proceedings of the 21th Annual Meeting of the Berkeley Linguistics Society*. S. 141-151.

- Hast, M. H. (1975): Experimental Physiology of the Larynx. In: Singh, S. (ed.) *Measurement Procedures in Speech, Hearing, and Language*. Baltimore: University Park Press. S. 345-362.
- Helfrich, H. (1979): Age markers in speech. In: Scherer, K. R., Giles, H. (eds.) *Social markers in speech*. Cambridge: Cambridge University Press. S. 63-107.
- Helfrich, H. (1985): *Satzmelodie und Sprachwahrnehmung*. Berlin: de Gruyter.
- Hermann, E. (1942): Probleme der Frage. Nachrichten von der Akademie der Wissenschaften in Göttingen. Philologisch-historische Klasse. Nr. 3-4.
- Hirst, D., Di Cristo, A. (1998): A survey of intonation systems. In: Hirst D., A. Di Cristo, A. (eds.) *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press. S. 1-44.
- Hombert, J.-M., Ohala, J., Ewan, W. (1979): Phonetic explanations for the development of tones. In: *Language* 55. S. 37-58.
- House, D. (1990): *Tonal Perception in Speech*. Lund: University Press.
- D'Imperio, M., House, D. (1997): Perception of Questions and Statements in Neapolitan Italian. In: *Proceedings of Eurospeech, vol. 1*. S. 251-254.
- Isačenko, A. V., Schädlich, H. J. (1970): *A Model of Standard German Intonation*. The Hague: Mouton.
- Jones, D., Ward, D. (1969): *The Phonetics of Russian*. Cambridge: University Press.
- Kirk, R. E. (1968): *Experimental design. Procedures for behavioral sciences*. Belmont (Cal.): Brooks/ Cole.
- Kirschbaum, E.-G. (2001): *Grammatik der russischen Sprache*. Berlin: Volk und Wissen Verlag.
- Klatt, D. H. (1973): Diskrimination of fundamental frequency in synthetic speech: implications for models of speech perception. In: *Journal of the Acoustical Society of America* 53. S. 8-17.
- Kodzasov, S. V. (1992): Review on Odé (1989) Russian intonation: a perceptual description. In: *Russian Linguistics. Vol. 16 (1)*. S. 146-155.
- Kodzasov, S. V. (1996): Kombinatornaja modelj frazovoj prosodii. [Ein kombinierendes Modell der Phrasenprosodie.] In: Nikolajeva T. M (ed.) *Prosoditsheskij stroj russkoj rietshi. [Prosodische Organisation der russischen Sprache.]* Moskva.
- Kodzasov, S. (1999): A new descriptive system for Russian intonation. In: *Proceedings of the XIV International Congress of Phonetic Sciences, Vol. 2*. San Francisco. S. 913-915.

Kohler, K. J. (1991a): Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. In: Kohler, K. J. (ed.) *Studies in German Intonation. Arbeitsberichte des Instituts für Phonetik Universität Kiel (AIPUK 25)*. S. 115-187.

Kohler, K. J. (1991b): A model of German Intonation. In: Kohler, K. J. (ed.) *Studies in German Intonation. Arbeitsberichte des Instituts für Phonetik Universität Kiel (AIPUK 25)*. S. 279-360.

Kohler, K. J. (1991c): Form and Function of Intonation Peaks in German: A Research Project. In: Kohler, K. J. (ed.) *Studies in German Intonation. Arbeitsberichte des Instituts für Phonetik Universität Kiel (AIPUK 25)*. S. 13-27.

Kohler, K. J. (1995): Einführung in die Phonetik des Deutschen. Berlin: Erich Schmidt Verlag.

Kohler, K. J. (2003): Neglected Categories in the Modelling of Prosody: Pitch Timing and Non-Pitch Accents. In: *Proceedings of the XV International Congress of Phonetic Sciences*. Barcelona. (im Druck)

Ladd, D. R. (1981): On intonational Universals. In: Myers, T. et al. (eds.) *The Cognitive Representation of Speech*. Amsterdam: North Holland Publishing. S. 389-399.

Ladd, D. R. (1996): Intonational phonology. Cambridge: Cambridge University Press.

Ladd, D. R., Silverman, K. E. A. (1984): Vowel intrinsic pitch in connected speech. In: *Phonetica* 41. S. 31-40.

Leed, R. L. (1965): A contrastive analysis of Russian and English intonational contours. In: *Slavic and East European Journal* 9. S. 62-75.

Lehiste, I. (1970): Suprasegmentals. Cambridge: The MIT Press

Leonard, R. J., Ringel, R. L. (1979): Vocal shadowing under conditions of normal and altered laryngeal sensation. In: *Journal of Speech and Hearing Research* 22. S. 794-817.

Lieberman, P. (1965): On the acoustic Basis of the Perception of Intonation by Linguists. In: *Word* 21, S. 40-54.

Lövqvist, A., Baer, T., McGarr, N. S., Story, R. (1989): The cricothyroid muscle in voicing control. In: *Journal of the Acoustical Society of America* 85 (3). S. 1314-1321.

Luukko-Vinchenzo, L. (1988): Formen von Fragen und Funktionen von Fragesätzen. Eine deutsch-finnische kontrastive Studie unter besonderer Berücksichtigung der Intonation. Tübingen: Max Niemeyer Verlag.

Makarova, V. (1999): Pitch Peak Alignment in Russian Declaratives, Interrogatives and Exclamations. In: *Proceedings of the XIV International Congress of Phonetic Sciences, Vol. 2*. San Francisco. S. 1173-1176.

McQueen, J. (1996): Phonetic Categorisation. In: *Language and Cognitive Processes* 11(6). S. 655-664.

Möbius, B. (1993): Ein quantitatives Modell der deutschen Intonation. Tübingen: Max Niemeyer.

Motsch, W., Viehweger, D. (1981): Sprachhandlung, Satz und Text. In: Rosengren, I. (ed.) *Sprache und Pragmatik: Lunder Symposium 1980*. Lund: Gleerup. S. 125-153.

Moulines, E., Carpentier, F. (1990): Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. In: *Speech Communication* 9, S. 453-467.

Niebuhr, O. (2003): Perceptual study of timing variables in F0 peaks. In: *Proceedings of the XV International Congress of Phonetic Sciences*. Barcelona. (im Druck)

Nikolajeva, T. M. (1977): Frasovaja intonatsija slavianskikh jazykov. [Phrasenintonation in slawischen Sprachen.] Moskva: Nauka.

Nikolajeva, T. M. (1989): Ob odnom skhodstvie slavianskoj i finnougorskoj frazovoj intonatsii. [Zu einer Ähnlichkeit zwischen slawischer und finno-ugrischer Phrasenintonation.] In: Bulatova, R.V., Nikolajeva, T.M. (eds.) *Slavianskoje i balkanskoje jazykoznanije: Prosodija*. [Slawische und balkanische Sprachwissenschaft: Prosodie.] Moskva: Nauka. S. 3-16.

Odé, C. (1987): A perceptual analysis of Russian intonation: an experiment. In: *Proceedings of the XI International Congress of Phonetic Sciences. Vol. 3*. Tallinn. S. 194-197.

Odé, C. (1989): Russian intonation: a percertual description. Amsterdam: Rodopi.

Odé, C. (1992): Perzeptivnaja ekvivalentnostj realizatsij tipov intonatsionnykh konstruksij E. A. Bryzgunovoj. [Perzeptive Äquivalenz der Realisierungen der Intonationskonstruktionen von E. A. Bryzgunova] In: *Studies in Russian Linguistics. Vol. 17*. Amsterdam: Rodopi. S. 227-284.

Ohala, J.J. (1978): Production of tone. In: Fromkin, V. A. (ed.) *Tone. A linguistic Survey*. London: Academic Press. S. 5-39.

Ohala, J. J. (1984): An Ethnological Perspective on Common Cross-Language Utilization of F0 of Voice. In: *Phonetica* 41. S. 1-16.

Ozhegov, S. U., Shviedova, N., J. (1993): Tolkovyj slovarj russkogo jazyka. [Erklärendes Wörterbuch der russischen Sprache.] Moskva.

Papoušek, M., Papoušek, H., Symmes, D. (1991): The meanings of melodies in motherese in tone and stress languages. In: *Infant Behavior and Development* 14. S. 415-440.

Paul, H. (1898): Prinzipien der Sprachgeschichte. Halle a. S.: Max Niemeyer.

- Peters, B. (1998): Prototypische Intonationsmuster in deutscher Lese- und Spontansprache. Magister-Arbeit, Universität Kiel.
- Pierrehumbert, J. (1980): The Phonology and Phonetics of English Intonation. Diss. MIT. Bloomington: Indiana University Linguistics Club.
- Pike, K. L. (1945): General Characteristics of Intonation. In: Bolinger, D. (ed.) (1972): *Intonation*. Harmondworth: Penguin. S. 53-82. (Auszug aus Pike, K. L. (1945): *The Intonation of American English*. Ann Arbor: University of Michigan Press. S. 20-41.)
- Pompino-Marschall, B. (1995): Einführung in die Phonetik. Berlin: De Gruyter.
- Reinhold Petersen, N. (1986): Perceptual compensation for segmentally conditioned fundamental frequency perturbation. In: *Phonetica* 43, 31-42.
- Repp, B. H. (1984): Categorical Perception, issues, methods, findings. In: Lass, N., J. (ed.) *Speech and Language: Advances in Basic Research and Practice. Vol. 10*. Academic Press: Orlando. S. 244-335.
- Rosental, D. E. (ed.) (1979): *Sovremennyy russkij jazyk*. [Moderne russische Sprache]. Moskva: Vysshaja shkola.
- Rossi, M. (1998): Intonation in Italian. In: Hirst D., A. Di Cristo, A. (eds.) *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press. S. 219-239.
- Saeed, J. I. (1997): *Semantics*. Oxford: Blackwell Publishers.
- de Saussure, F. (1931): *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin/Leipzig.
- Savtshienko, A. N. (ed.) (1962): *Orfografija i punktuatsija russkogo jazyka*. [Russische Orthographie und Interpunktion.] Rostov.
- Schiefer, L. Batliner, A. (1988): Intonation, Ordnungseffekt und das Paradigma der kategorialen Wahrnehmung. In: Altmann, H. (ed.) *Intonationsforschungen*. Tübingen: Max Niemeyer Verlag.
- Shviedova, N. J. (ed.) (1970): *Grammatika sovremennogo russkogo literaturnogo jazyka*. [Grammatik der modernen russischen Standardsprache]. Moskva: Nauka.
- Sundberg, J. (1979): Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7. S. 71-79.
- Svetozarova, N. D. (1982): *Intonatsionnaja sistema russkogo jazyka*. [Intonationssystem der russischen Sprache.] Leningrad.

Svetozarova, N. D. (1998): Intonation in Russian. In: Hirst D., A. Di Cristo, A. (eds.) *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press. S. 261-275.

Takefuta, Y. (1975): Method of Acoustic Analysis of Intonation. In: Singh, S. (ed.) *Measurement Procedures in Speech, Hearing, and Language*. Baltimore: University Park Press. S. 363-378.

Thon, W. (1982): Microprocessor-controlled reaction-time measurement. In: Barry, J., Kohler, K. J. (eds.) *Arbeitsberichte des Instituts für Phonetik Universität Kiel (AIPUK 18)*. S. 84-91.

Thorsen, N. (1979): Interpreting Raw Fundamental-Frequency Tracing of Danish. In: *Phonetica* 36. S. 57-78.

Thorsen, N. (1985): Intonation and text in Standard Danish. In: *Journal of Acoustical Society of America* 77 (3). S. 1205-1216.

Torsujeva, I. G. (1979): Intonatsija i smysl vyskazyvanija. [Intonation und Bedeutung einer Äußerung.] Moskva: Nauka.

Wenk, R. (1975): Die Intonation. In: Gabka, K., Wiede, E. (eds.) *Einführung in das Studium der russischen Sprache. Bd. 1: Phonetik und Phonologie*. Leipzig: VEB. S. 175-198.

Werner, S., Keller, E. (1994): Prosodic aspects of speech. In: Keller, E. (ed.) *Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges*. Chichester: John Wiley. S. 23-40.

Whalen, D., Levitt, A. (1995): The universality of intrinsic F0 of vowels. In: *Journal of Phonetics* 23. S. 349-366.

Williamson, K. (1979): Sentence tone in some southern nigerian languages. In: *Proceedings of the IX International Congress of Phonetic Sciences. Vol. 2*. S. 424-430.

Wunderlich, D. (1976): Studien zur Sprechakttheorie. Frankfurt am Main: Suhrkamp.

Yuan, J., Shih, Ch., Kochanski, G. P. (2002): Comparison of Declarative and Interrogative Intonation in Chinese. In: *Proceedings Speech Prosody 2002*. Aix-en-Provence. URL4.

Zimmer, D. E. (1988): So kommt der Mensch zur Sprache. Zürich: Haffmans.

URL1: <http://www.ipds.uni-kiel.de/forschung/xassp.de.html>

URL2: <http://www.fon.hum.uva.nl/praat>

URL3: <http://www.lpl.univ-aix.fr/sp2002/pdf/gussenhoven.pdf>

URL4: <http://www.lpl.univ-aix.fr/sp2002/pdf/yan-shih-kochanski.pdf>



## Inhaltsverzeichnis

<b>1. Einleitung.....</b>	<b>1</b>
<b>2. Theoretische Grundlagen .....</b>	<b>4</b>
2.1. Problemfeld der Intonation .....	4
2.1.1. Intonation als kommunikatives Mittel.....	4
2.1.2. Manifestationsebenen der Intonation .....	6
2.1.3. Physiologische Bedingungen.....	7
2.1.4. Intonationskategorien .....	9
2.1.5. Methoden der Intonationsforschung.....	12
2.2. Sprachliche Kategorien der Frage vs. Aussage.....	15
2.2.1. Semantik und Pragmatik.....	15
2.2.2. Formale Kodierung der Frage im Sprachvergleich .....	16
<b>3. Charakteristika der russischen Intonationsmuster aufgrund der akustischen Analyse von Produktionsdaten.....</b>	<b>21</b>
3.1. Beschreibungen der russischen Intonation in der Literatur .....	21
3.2. Ergänzende Voruntersuchung .....	30
3.2.1. Sprachmaterial und Methode.....	30
3.2.2. Ergebnisse.....	31
3.2.3. Diskussion .....	33
<b>4. Ansätze der Beschreibung von Intonationskontrasten in der Perzeption .....</b>	<b>36</b>
4.1. Perzeptionsuntersuchungen zum Intonationssystem des Russischen .....	36
4.2. Perzeptionsuntersuchungen zur intonatorischen Satzmodusunterscheidung in anderen Sprachen.....	40
4.2.1. Gósy und Terken (1994) zum Ungarischen.....	40
4.2.2. D`Imperio und House (1997) über die neapolitanische Varietät des Italienischen.....	42
4.3. Evidenzen für das Russische.....	44
<b>5. Empirische Untersuchung zur Perzeption der Frage im Russischen.....</b>	<b>45</b>
5.1. Hypothesen .....	45
5.2. Methode .....	46
5.2.1. Sprachmaterial .....	46
5.2.2. Anfertigung der Experimentstimuli.....	48
5.2.2.1. Stilisierung der Originalkonturen .....	48

5.2.2.2. Manipulationsdesign .....	51
5.2.2.3. Herstellung der Stimulusreihen.....	52
5.2.2.4. Anmerkungen zu den durchgeführten Manipulationen .....	56
5.2.3. Vorbereitung der Experimente .....	57
5.2.4. Versuchspersonen.....	59
5.2.5. Experimentdurchführung.....	59
5.3. Ergebnisse .....	60
5.3.1. Besonderheiten der Stichprobe.....	60
5.3.2. Versuchsplan und Auswertung.....	61
5.3.3. Ergebnisse der Identifikationstests .....	64
5.3.4. Ergebnisse der Diskriminationstests.....	68
5.3.5. Ergebnisse der Vortests .....	73
<b>6. Diskussion .....</b>	<b>78</b>
6.1. Interpretation der Ergebnisse .....	78
6.2. Universale Tendenzen im Sprachvergleich .....	82
6.3. Kritische Bemerkungen zum durchgeführten Experiment .....	86
6.4. Vorschläge für weiterführende Forschung.....	87
<b>7. Zusammenfassung und Ausblick.....</b>	<b>89</b>
<b>Literaturliste.....</b>	<b>91</b>