

**Bemerkungen zu Gussenhoven**  
**"Experimental approaches to establishing gradience vs discreteness in intonation"**  
**(Experimental Prosody Research Workshop, Leipzig, 8.10.04)**

*K. J. Kohler, IPDS, Kiel*

24.10.2004

1. Grundlegende Frage:  
Ist der Unterschied zwischen zwei Äußerungen phonologisch oder phonetisch?  
Sie basiert auf der für alles Sprachliche postulierten Aufteilung in "cognitive representations (phonology)" und "the variation in their articulation (phonetic implementation/pronunciation)". Damit haben wir die alte Dichotomie Phonologie – Phonetik, die ontologischen Status für Sprache und Sprechen bekommt. Aus der Dichotomie folgt, daß die kognitive Repräsentation ein diskretes mentales Objekt ist, die artikulatorische Variation hingegen eine unendliche Klasse physischer Manifestationen.
2. Wie kann im Bereich der Intonation die Frage entschieden werden, ob ein Äußerungsphänomen "phonologisch" oder "phonetisch" IST? Die Entscheidung ist Voraussetzung für die Erstellung des phonologischen Inventars und zum Verständnis der Grammatik.
3. Tonhöhen/Frequenzvariabilität ist paralinguistisch bedeutungsvoll, aber darunter verbergen sich die eigentlichen kognitiven linguistischen Strukturen, die es zu erfassen gilt. Damit haben wir die zweite ontologisierte Dichotomie zwischen "linguistisch" und "paralinguistisch". Folglich kann ein Rückgriff auf unterschiedliche Oberflächenbedeutungen zweier Äußerungen nicht entscheiden, ob sie verschiedenen oder einer kognitiven linguistischen Repräsentation zuzuordnen sind. Die Verwendung von "kognitiv" in diesem Zusammenhang macht gleichzeitig deutlich, welches Element in den Paarungen "phonologisch/phonetisch" und "linguistisch/paralinguistisch" Priorität zugewiesen bekommt.
4. Linguistisch sind die diskreten Kategorien "pitch accent" (vorhanden oder nicht) und "melodische Einheiten" (Oppositionen von L und H bzw. deren Sequenzen). Sie sind als existent gegeben und werden nicht hinterfragt. Im Gegenteil, die phonetische Variabilität muß auf sie ohne Rest projiziert werden. Andererseits sind H-Maximum und L-Minimum sowie f<sub>0</sub>-Alignment graduell skaliert und deshalb nicht linguistisch. Falls sie zur Semantik der Äußerung beitragen, sind sie folglich paralinguistisch. Diese Ansicht findet sich bei der Schülerin Aoju Chen in ihrem Leipziger Beitrag in allen Details wieder.
5. Hierbei zeigt sich gleichzeitig, welche Bedeutung der Prüfstatistik in der Beantwortung dieser linguistischen Fragen zugeschrieben wird. Das Paradigma der kategorialen Wahrnehmung, das in einigen referierten Untersuchungen zentral war, liefert weitgehend keine oder keine klaren Diskriminationsmaxima bei Kategorienwechsel in der Identifikation. Damit scheidet es als Instrument der phonologischen, d.h. linguistischen Musterbestimmung aus. Selbst die Anwendung des Paradigmas in der Kieler Gipfelverschiebung, die ja Kategorialität für frühe vs. mittlere Gipfel nachgewiesen hat, wird ausgeklammert und dem Alignment zugerechnet. Daß hier völlig verschiedene Phänomene unter einem konzeptuellen Dach zusammengefaßt werden, hängt mit zwei Voraussetzungen der AM-Phonologie zusammen:

- (a) die Semantik der Gipfelverschiebung wird im paralinguistischen Bereich angesiedelt,
- (b) die ToBI-artige Notation mit H und L erfaßt nicht das unterschiedliche Timing, sondern postuliert unterschiedliche melodische Muster (z.B. HL\* vs LH\*); diese sind aber mit dem Paradigma der kategorialen Wahrnehmung nicht erfaßbar.
6. So verweist der Autor dann schließlich auf die Untersuchung Gussenhoven/Rietveld 'The behavior of H\* and L\* under variations in pitch range in Dutch rising contours', Es wird ein Verfahren vorgestellt, das von den unterschiedlichen Intonationsmustern "high rise" und "low rise" sowie ihren Tonakzentnotationen H\*H-H% und L\*H-H% ausgeht und daran die Frage nach ihrer phonologischen bzw. phonetischen Differenzierung knüpft. Wenn gezeigt werden kann, daß sich die mit H\* und L\* erfaßten abweichenden Stellen in diesen Mustern unter eindeutig gradueller Variation unterschiedlich verhalten, dann müssen die Töne verschiedenen Kategorien angehören. Es wurde 'pitch range' an den Stellen H\*/L\* und H% variiert, d.h. ein Parameter, dem graduelle Variabilität zugeschrieben wird. Hörer mußten dann die resultierenden Konturen hinsichtlich eines expressiven (paralinguistischen) Parameters, nämlich "surprise", skalieren, von dem angenommen werden kann, daß er mit der graduellen Veränderung von 'pitch range' kovariiert. Das Ergebnis zeigt ein stärkeres 'surprise'-Urteil für höheres H einerseits und für tieferes L\* andererseits. Das wird in Beziehung zu der Untersuchung von Liberman-Pierrehumbert gesetzt, die eine Korrelation zwischen 'pitch range' und H-Erhöhung bzw. L\*-Absenkung für Fokus festgestellt haben. Damit ist bewiesen, daß die beiden Tonmuster kategorisch verschiedene phonologische Strukturen repräsentieren und daß diese tatsächlich H\* und L\* sind.
7. Gegen diese Argumentation sind folgende Punkte anzuführen
- Der Status der beiden Notationen wird nicht klar definiert. An sich muß es sich um eine abstrahierende, aber dennoch *phonetische* Beschreibung handeln; denn der phonologische Status soll ja erst eruiert werden. Aber in dieser Notation werden bereits die Kategorien von 'pitch accents' und 'boundary tones' der *AM-Phonologie* verwendet. Damit befinden wir uns in einem Argumentationszirkel.
  - Dieser bereits am Ausgangspunkt der Untersuchung vorhandene phonologische Ansatz bedingt dann die punktuelle Variation im Muster und den Bezug der Ergebnisse der globalen Stimulusbeurteilung durch die Hörer auf genau diese tonalen Punkte.
  - Die Synthese der beiden Muster in nicht-finaler Satzposition generiert einen Knick im Anstieg mit einem deutlichen Plateau; das führt zu einem Nebenakzent auf dem Wort, mit dem der Aufstieg fortgesetzt wird, d.h. 1 in PROLAB-Notation. Das wird jedoch in AM/ToBI ignoriert, hat aber bei einem Experiment wie dem vorliegenden Konsequenzen für die (globale) Höhen/Tiefen-Beurteilung, folglich für 'pitch range' und für die Stärke des 'surprise'-Urteils. So zeigt L\* in dieser Position kaum noch den abnehmenden Effekt mit zunehmender Höhe von L\*, und außerdem hat das Muster generell durchgehend die höchsten 'surprise'-Werte. Umgekehrt hat H\* hier die niedrigsten 'surprise'-Werte. Diese Ergebnisse hängen sicher damit zusammen, daß die globalen Konturen für H\* und L\* final und nicht-final unterschiedlich sind.
  - Syntaktisch markierte und unmarkierte Entscheidungsfragesätze wurden gemischt. 'surprise' manifestiert sich aber unterschiedlich in diesen. Außerdem ist

'surprise' in H\*H-H% und L\*H-H% nicht identisch. Und schließlich spielt die kontextuelle Einbettung eine wichtige Rolle für die Zuordnung einer Äußerung zu einem Grad von 'surprise'.

Die folgenden Dialogbeispiele illustrieren diese Einflußfaktoren auf das Urteil 'surprise'.

**A. Wo ist er denn eigentlich?**

**B. Er ist in Rom. (Ist er in Rom?)**

**H\*H-H%** ist hier genauso wenig möglich wie die syntaktische Fragekonstruktion (mit beiden Mustern), weil nach der Informationsfrage des Sprechers A keine Frage des Sprechers B an A stehen kann, weder eine Rückfrage noch eine ja-nein-Entscheidungsfrage. Eine Entscheidungsfrage ist in dieser Situation nur möglich, wenn durch A eine Frage an eine Gruppe gestellt wird und B mit einer Entscheidungsfrage zu einer Informationshypothese an die Gruppe nachhakt.

**L\*H-H%** ist möglich und bedeutet eine persönliche Bewertung der kontrastierenden Darstellung eines Sachverhalts durch den Sprecher: "Du weißt das nicht! Das ist doch längst bekannt!" Entscheidend für die Signalisierung dieser Semantik ist der Kontrast zwischen tief und hoch in der akzentuierten Silbe; je tiefer der Vorlauf vor dem Aufstieg, desto stärker diese persönliche negative Bewertung. Das ist 'surprise' im Sinne von persönlich bewerteter Abweichung.

**A. Er ist nach Rom gefahren.**

**B. Er ist in Rom?**

**H\*H-H%** ist möglich und bedeutet "ist das tatsächlich so?" = erstaunte Nachfrage, ohne Bewertung.

**L\*H-H%** das ist die Nachfrage mit persönlicher Bewertung.

Die Beurteilung auf einer einzigen Skala 'surprise' kann diesen Unterschied nicht erfassen.

**Ist er in Rom?** mit H\*H-H% ist komisch, weil das syntaktische Muster die Nachfrage unwahrscheinlich macht, aber nur die Nachfrage Sinn ergibt. Die Ortsinformation ist durch Sprecher A bereits geliefert, eine ja-nein-Entscheidung darüber kann sinnvollerweise von A nicht mehr verlangt werden, nur noch die Bestätigung einer ungläubigen Nachfrage. Hätte A gesagt "Er ist nach Italien gefahren.", dann wäre "Ist er in Rom?" mit H\*H-H% möglich. Diese spezifische Ortsinformation ist noch nicht gegeben, über sie wird eine Entscheidung verlangt. In dieser Situation wäre nun andererseits "Er ist in Rom?", d.h. die ungläubige Nachfrage unwahrscheinlich.

**Ist er in Rom?** mit L\*H-H% ist möglich als Nachfrage mit persönlicher Bewertung, die sich auf das "in Rom sein" bezieht. **Er ist in Rom?** mit L\*H-H% bezieht sich als persönlich bewertete Nachfrage allein auf "Rom".

**A. Ich hab 100 Euro fürs Rote Kreuz gespendet.**

**B. 100 Euro ist nicht viel.**

**H\*H-H%** ist unwahrscheinlich: A hat kein Urteil über die Größe des Geldbetrags abgegeben, deshalb macht eine Nachfrage, ob das tatsächlich nicht viel ist, keinen Sinn.

Mit **L\*H-H%** verneint B, daß es sich um einen großen Betrag handelt, und bewertet dies im Gegensatz zur Aussage von A: "Das ist aber nicht viel."

**A. Ich hab 100 Euro fürs Rote Kreuz gespendet.**

**B. Ist 100 Euro nicht viel?**

**H\*H-H%** ist möglich als distanzierte Frage, ob das nicht zu viel ist. Mit dieser Frage wird nicht verneint, daß der Betrag groß ist, sondern gerade im Gegenteil B zur Kommentierung aufgefordert, daß dem so ist.

**L\*H-H%** ist ebenfalls möglich, diesmal mit persönlicher Bewertung der erfragten Kommentierung: "ich meine, das ist doch etwas viel".

**A. Ich hab nicht viel gespendet. Nur 100 Euro.**

**B. 100 Euro ist nicht viel?**

**H\*H-H%** ist möglich als Nachfrage, ob A wirklich meint, daß 100 Euro nicht viel ist.

**L\*H-H%** scheint nicht zu passen, weil sich weiterhin die Interpretation aufdrängt, daß B eine persönlich bewertete Aussage "ist nicht viel" macht, anstatt die Verneinung durch A in einer Rückfrage persönlich zu bewerten. Das wird besser durch die Konstruktion der Entscheidungsfrage ausgedrückt, und dann am besten durch Akzentuierung von "nicht".

**A. Ich hab nicht viel gespendet. Nur 100 Euro.**

**B. Ist 100 Euro nicht viel?**

**H\*H-H%** signalisiert eine Routinefrage, ob 100 Euro nicht doch viel ist. Die ausgewählte Äußerung kann in der Perzeption zwischen Akzentuierung auf "nicht" und "viel" springen. In dem Kontext "nicht viel gespendet" wird der identische Stimulus eher auf "nicht", im anderen Fall eher auf "viel" akzentuiert wahrgenommen. Ersteres ist dann auch nicht mehr **H\*H-H%**, sondern **L\*H-H%**.

**L\*H-H%** signalisiert eine Bewertung der Frage durch den Sprecher: "Das ist doch viel!"

- Diese Beispiele zeigen, daß die Skalierung von kontextfreien Einzelsätzen in Hörexperimenten ein falsches Bild der Substanz-Funktionsbeziehung liefert, noch dazu wenn syntaktisch unterschiedliche Sätze gemischt und anschließend zur statistischen Bearbeitung zusammengefaßt werden. Wir müssen zu kontextualisierten Hörexperimenten kommen.
- Eine einzige Bewertungsskala vermittelt dann erst recht ein falsches Bild, wenn die Vpp gezwungen werden, verschiedene semantische Aspekte der Konturen auf einer Skala abzubilden.
- Das entscheidende semantische Merkmal von **L\*H-H%** ist die kontrastierende Bewertung eines Sachverhalts durch den Sprecher. Dieses semantische Merkmal wird durch den tief-hoch-Kontrast im akzentuierten Vokal getriggert. Es fehlt in **H\*H-H%**, wo durch die Verstärkung der Höhen ab der akzentuierten Silbe bis zum Ende der Äußerung eher eine erstaunte Nachfrage signalisiert wird. Damit ist nicht ein punktueller Ton entscheidend, sondern die Gestaltung eines globalen Musters.
- Die Anhebung von **H\***, aber Absenkung von **L\*** zur Erzielung einer semantischen Verstärkung des Erstaunens bzw. der persönlichen kontrastierenden Bewertung ist damit völlig anders zu sehen als die Ausprägung von 'pitch range'

in Gipfelkonturen zur Fokussierung bei Liberman-Pierrehumbert. Dort ging es einheitlich um dasselbe Phänomen: Fokus; hier geht es um unterschiedliche Phänomene, die durch die einheitliche Bewertungsskala 'surprise' verdeckt werden.

- Damit entfällt die phonologische Argumentation: dieses Experiment beweist nicht den phonologischen Status als H\* und L\*.
  - Es ist aber auch klar geworden, daß die Trennung in 'linguistisch' und 'paralinguistisch' der Erfassung intonatorischer Phänomene abträglich ist.
8. Es könnte eingewandt werden, daß meine Argumentation vom Deutschen ausgeht, während Gussenhoven/Rietveld holländische Daten diskutierten. Wir können jedoch davon ausgehen, daß die Sprachen weitgehend konform gehen, auch in diesen Punkten. Im übrigen weisen die Autoren selbst darauf hin: sie ziehen das Englische heran, wo die Muster differenziert werden, und sprechen die Nähe der westgermanischen Sprachen zueinander an.
9. In KIM sind die Muster unterschieden als frühes bzw. spätes Tal mit <?>.