

Bemerkungen zu Atterer/Ladd

"On the phonetics and phonology of "segmental anchoring" of F0, evidence from German"

K. J. Kohler, IPDS, Kiel

24.10.2004

1. Die AM-Phonologie setzte für die Intonationsbeschreibung 'pitch accents' und 'boundary tones' an, die entweder H oder L sein können. Diese Töne können sich im Fall der 'pitch accents' auch zu einigen weiteren bitonalen Akzenten verbinden. Alle diese Einheiten sind phonologisch, d.h. Einheiten der linguistischen Form, die zwischen linguistischer Bedeutung und phonetischer Substanz vermitteln. Eine Direktbeziehung zwischen Bedeutung und Substanz wird ebenso ausgeschlossen, wie paralinguistische Bedeutung aus dem Untersuchungsfeld ausgeklammert wird. (Pierrehumbert: association)
2. Die phonologischen Einheiten werden Silben, d.h. der Artikulation zugeordnet. Im Fall der bitonalen Einheiten erhebt sich die Frage, welcher der beiden Töne, der erste oder der zweite, primär mit der Akzentsilbe assoziiert ist und welcher "leading" oder "trailing" ist. Diese Frage wurde durch die Einführung des * entschieden, und so wurde z.B. zwischen L+H* und L*+H unterschieden. (Pierrehumbert)
3. Als man nun im Geiste der Laborphonologie die Einheiten der Intonationsphonologie ins Labor trug, um sie mit Meßsubstanz zu füllen, stellte man fest, daß z.B. in L+H* der primär mit der Akzentsilbe assoziierte H-Ton das F0-Maximum gar nicht immer in der Akzentsilbe erreichte, sondern erst in der nachfolgenden unakzentuierten Silbe. Man behielt aber den phonologischen Ansatz der Töne unverändert bei und führte nun die sekundäre Zuordnung von L und H in bitonalen Einheiten zu unterschiedlichen Stellen in der Silbenstruktur ein. Das war wiederum eine phonologische Abstraktion von der Zeitachse. (Pierrehumbert & Beckman: secondary association)
4. Die laborphonologische Messung der phonologischen Töne ging aber weiter, und man fragte nun nach der zeitlichen Verankerung der Tonakzente, insbesondere der bitonalen, in der Silbenstruktur. Damit begann sich die Messung von dem phonologischen Tonansatz zu lösen, die Messung entwickelte ein Eigenleben. (Untersuchungen zum alignment, u.a. Ladd et al)
5. Schließlich kehrte sich das Verhältnis von phonologischer Form zu phonetischer Substanz um. Die Messung wurde primär und führte zu einem "phonologischen" Ansatz LH oder HL. Damit wird aber nicht nur das linguistische Axiom von Form und Substanz aufgegeben, sondern auch der Bezug zu sprachlichen Einheiten und ihrer Funktion. Es wird gar nicht mehr danach gefragt, welche intonatorischen Muster in einer Sprache (oder beim Sprachvergleich in jeder der verglichenen Sprachen) zur funktionalen Differenzierung eingesetzt werden. Sprecher bekommen konstruierte Sätze zum Lesen vorgelegt, und alles, was in den Daten z.B. unter eine phonetische Interpretation von LH fällt, wird hinsichtlich F0 (und zugehöriger Zeitparameter) gemessen. Darüber werden Mittelwerte und Standardabweichungen berechnet und über ANOVAs prüfstatistische Aussagen abgeleitet. Prüfstatistische Signifikanzen werden dann als phonologisch-prosodische Kategorisierungen interpretiert. (Atterer/Ladd)
6. Damit ist die Phonetik ins 19. Jh. zurückgekehrt. Das ist exakt die Wissenschaftsphilosophie der frühen Experimentalphonetik eines Rousselot, eines Scripture, eines Panconcelli-Calzia. Alle sprachlichen Kategorien lösen sich auf, die Wahrheit liegt in den Kurven und den Zahlen, die daraus abgeleitet werden. Das ist die Einstellung einer bestimmten Gruppe "moderner Speech Scientists", die entweder selbst Sprachtechnologe sind oder deren Denken durch die Sprachtechnologie bestimmt wird.

7. Damit gerät aber die grundlegende Aufgabe der phonetischen Forschung aus dem Blickfeld. Lautsprache ist ein kultursoziologisches Phänomen menschlicher Interaktion, das über physiologische und physikalische Träger funktioniert. Damit ist die fundamentale Fragestellung lautsprachlicher Forschung "Wie funktioniert lautsprachliche Kommunikation in Situationen menschlicher Interaktion, d.h. welches sind die Beziehungen zwischen kommunikativen Funktionen und phonetischen Signalen für Sprecher und Hörer in situativer Einbettung?" Damit steht zwar nicht die linguistische Form im Vordergrund, aber erst recht nicht die Signalmessung, sondern die kommunikative Funktion, der Signale zugeordnet werden. Wer das nicht erkennt und befolgt, mag sich vielleicht sehr modern geben, in einer praktischen Ingenieursicht, ist aber in wissenschaftstheoretischer Hinsicht ein Steinzeitmensch.

8. Nun ein paar Details der Arbeit von Atterer/Ladd, die das bereits Gesagte noch unstreichen und gleichzeitig die Schwäche der angewandten Methode aufzeigen.

- Es wird hier eine Differenzierung in zwei Sprechergruppen "Norddeutsch" und "Süddeutsch" nach geographischen Gesichtspunkten durchgeführt. Die norddeutschen Sprecher haben die meiste Zeit ihres Lebens in Norddeutschland verbracht, die süddeutschen in Bayern. Damit wird zum einen nicht differenziert zwischen dem südwestdeutschen (südrheinfränkischen, alemannischen) und dem südostdeutschen (nieder-/oberbayrischen, oberfränkischen) Sprachgebiet, die beide wiederum untergliedert sind und sehr unterschiedliche Ausprägung zeigen, gerade was Intonation anlangt. Zum andern wird aber auch gar nicht unabhängig eruiert, ob die bayrischen Probanden tatsächlich alle bayrische Züge in ihrer Aussprache hatten; die Tatsache, daß sie ihr Leben lang in Bayern gelebt haben, bedeutet nicht, daß sie auch bayrisch klingen. In München leben sehr viele Nordlichter. Und wenn Bewohner Bayerns bayrisch klingen, kann das in sehr unterschiedlichem Grad sein, was wiederum insbesondere für eine Intonationsuntersuchung von grundlegender Relevanz ist. Daß diese Bedenken hinsichtlich der Validität der Untersuchungsbasis für die Fragestellung berechtigt ist, geht aus folgendem Zitat hervor: "... a majority of the speakers can readily be assigned to the correct group on a few seconds' impressionistic listening..." Also nicht alle.

- Fast ein Viertel der aufgenommenen Probanden wurde eliminiert, u.a. wegen "reading too fast and monotonously (which would make identifications of F0 minima and maxima unreliable)..." Das ist methodisch unsauber; denn die ausgeschiedenen Sprecher waren ja auch süddeutsch oder norddeutsch, wenn sie auch nicht die erwarteten LH-Konturen produzierten. Die Schlußfolgerungen aus den Daten der verbliebenen Sprecher beziehen sich aber auf das Nord- bzw. Süddeutsche schlechthin.

- Der nächste methodische Kritikpunkt zielt in eine ähnliche Richtung. Auch einzelne Realisierungen der schriftlich vorgegebenen Sätze wurden ausgeklammert: "Occasionally, test sentences were discarded because the speaker used a completely different intonation contour from the one under investigation." Hier kommt noch hinzu, daß "completely different" überhaupt nicht beschrieben wird, d.h. das Ausschlußkriterium bleibt völlig unklar.

- Es wird auch nichts darüber gesagt, wie die Daten gelesen wurden. Der Hinweis "The list of German sentences read by the speakers ..." sowie der Bezug auf Ladd et al 1999 läßt vermuten, daß die Liste abgelesen wurde. Damit kommt aber eine Listenintonation herein. Die Datenerhebung kann also eine weitere intervenierende Variable enthalten, die für die Ausprägung der Intonation wichtig ist. Außerdem werden keine Angaben über Wiederholungen gemacht; es kann nur angenommen werden, daß die Liste von jeder Versuchsperson einmal gelesen wurde.

- Auf der anderen Seite wird hervorgehoben, daß sich Sprecher verschiedener Sprachen in irgend einer Weise ähnlich verhalten: "...the evidence presented here suggests that the ordinary accentual rise used spontaneously by speakers of English, Dutch, German and Greek on the first accented word of read sentences is in some way the "same thing" cross-linguistically." "in some way" ist völlig undefiniert, und "the same thing" ebenfalls. Was für ein "Ding" ist das, worum es hier geht? Und wann wird dieses "Ding" gleich? Dieses "Ding" ist eine steigende Intonation (was damit nicht konform ging, wurde ausgeschlossen). Dieses LH wird vorweg postuliert und dann ausgemessen. Aufgrund der Ergebnisse wird dann gesagt, daß es sich trotz phonetischer Unterschiede irgendwie um dasselbe "Ding" handeln muß. Das ist ein Zirkelschluß, der überhaupt keine Aussagekraft hat.

- Die Basisarbeit für die englischen Daten englischer Sprecher (Ladd et al 1999) macht es nun sehr wahrscheinlich, daß die englischen Probanden zumindest zum Teil andere F0-Muster verwendeten als die deutschen. Die Abbildung für "minimal fee" zeigt eine klare Gipfelkontur mit Einbuchtung, die Abbildung für "(in) Ermangelung (eines Lehrers)" den ersten Teil eines Hutmusters, d.h. im deutschen Beispiel haben wir im Gegensatz zum englischen ein Auslaufen auf einem Plateau und den Abstieg erst auf der nächsten Akzentsilbe (im nächsten Wort), im englischen Beispiel haben beide Akzentstellen eigene Gipfelkonturen mit An- und Abstieg. Die beiden steigenden Konturen sind also unterschiedlich in den Gesamtverlauf eingebunden, und das ist auch für die Ausprägung des Anstiegs relevant. Wenn der Abstieg im ersten akzentuierten Wort bereits erforderlich ist, muß der Anstieg so erfolgen, daß dies möglich wird. Damit ist die Wahrscheinlichkeit groß, daß in der Gipfelreihe des Englischen der erste Anstieg etwas früher erfolgt als der erste Anstieg im Hutmuster des Deutschen.

- Wenn nun des weiteren berücksichtigt wird, daß die englischen Sätze überwiegend die Konstruktion Adjektiv + Substantiv hatten, dann ist hierfür die Gipfelreihe üblicherweise zu erwarten. Im Deutschen haben aber 9 der 13 Sätze die Konstruktion Substantiv + Genitivattribut, wo das Hutmuster typisch ist. Es hätte also vor der Messung untersucht werden müssen, ob in den englischen bzw. deutschen Daten die gleichen Hut- bzw. Gipfelreihenmuster produziert wurden. Diese Forderung zielt zum einen darauf, daß Sätze nicht nur nach phonetischen und phonologischen, sondern auch nach funktionalen Gesichtspunkten zusammengestellt werden müssen. Zum andern weist sie darauf hin, daß Intonation nicht aus Gesamtverläufen herausgerissen werden darf.

- Ferner liegt in 7 der 15 englischen Sätze der hohe Vokal /I/ vor. Dieser kommt unter den 13 deutschen Sätzen nur einmal vor. Damit ist eine mikroprosodische Bias in Richtung auf ein höheres F0 im Vokal nicht auszuschließen, d.h. ein früheres Erreichen des Maximums.

- Wenn nun noch hinzugenommen wird, daß die englischen Daten nicht die gleiche Sprechgeschwindigkeit repräsentieren wie die deutschen, ist die Vergleichsbasis für die Daten dünn. Wenn ferner berücksichtigt wird, daß die Unterschiede in der Positionierung des Maximums recht klein sind und in den prüfstatistischen Vergleichen der verschiedenen Stichproben nicht einmal immer signifikant sind, ist schwer nachzuvollziehen, daß es sich hier um eine progressive Rechtsverschiebung des Maximums in der Ordnung Süddeutsch > Norddeutsch > Englisch, Holländisch handelt, die diese Sprachen als ganze charakterisiert. Die Verschiebung für L ist zwar stärker ausgeprägt und immer signifikant, aber damit ist noch nichts darüber gesagt, wie der schwache H- und der stärkere L-Effekt *zusammen* interpretiert werden müssen. Das kann auf dem Wege der Signalmessung und statistischen Prüfung nicht entschieden werden.

- Schließlich ist bereits von den Gutachtern angemerkt worden, daß das Umfeld intonatorischer Oppositionen in den verglichenen Sprachen berücksichtigt werden müßte. Wieviele durch

unterschiedliche Synchronisierung differenzierte LH-Muster müssen in jeder Sprache angesetzt werden? Für das Deutsche brauchen wir mindestens zwei, vermutlich drei; für das Englische gilt dasselbe. Dann ist die Frage, wie sprachliche Unterschiede in früherem oder späterem Alignment noch in dieses Oppositionssystem eingebaut werden können. Werden sämtliche distinktiven LH-Synchronisierungen zur unterschiedlichen phonetischen Ausprägung in den Sprachen verschoben? Oder erfaßt die Erhebungsmethode unterschiedliche Glieder der Oppositionskette für die verschiedenen Sprachen? Oder verwenden die Sprecher dieser Sprachen die unterschiedlichen Glieder mit unterschiedlicher Häufigkeit und produzieren deshalb unterschiedliche Prototypen beim prototypischen Lesen?

- Wenn alle diese Punkte zusammengenommen werden, lassen sich die Schlußfolgerungen der Autoren aus ihrer Datenerhebung und –analyse nicht nachvollziehen. Sie haben nicht das nachweisen können, was sie vorgeben gezeigt zu haben.

- Selbst wenn sie tatsächlich eine nachvollziehbare Beweisführung an ihren akustischen Produktionsdaten geliefert hätten, bliebe immer noch die Frage nach der perzeptorischen Relevanz der Verschiebungen von L und H. Was bedeuten die Meßunterschiede für die Wahrnehmung? Ist daran der Dialektunterschied unter anderem festmachbar?

9. Vor diesem Hintergrund ist der abschließende Paragraph geradezu verwegen. Hier liest man die gleichen Worte wie aus der Feder der ersten Experimentalphonetiker. Feine instrumentelle Messung ersetzt nicht nur bisherige phonetische Untersuchung, sondern ist auch infinit besser, weil letztere die in der instrumentellen Messung enthaltene Wahrheit gar nicht erfassen kann.

10. Das ist eine wissenschaftliche Einstellung, die sich zunehmend verbreitet, ausgehend von signalverarbeitenden Ingenieuren, die zwar die Signalanalyse beherrschen, aber von Sprache nicht mehr als eine rudimentäre Ahnung haben. Statt Ingenieuren die Fragestellungen lautsprachlicher Grundlagenforschung, d.h. jenseits technologischer Anwendungen, näherzubringen, übernehmen "moderne Speech Scientists" mit den technologischen Werkzeugen auch die Sichtweise und prägen damit phonetische Handbücher und phonetische Zeitschriften, wie z.B. Journal of Phonetics, sowie phonetische Studiengänge, die über das Erlernen von Verfahren der Sprachverarbeitung kaum hinausgehen und nicht mehr hinführen zu einer Erfassung von Zusammenhängen lautsprachlicher Kommunikation in soziokulturellen Kontexten. Das ist eine wissenschaftliche Revolution im Kuhnschen Sinne, die zwangsläufig zu einer Gegenrevolution führen muß, sobald die Richtungszwänge eklatant werden. Dabei lernt jede Revolution aus der vorangehenden, und zwar sowohl aus ihren negativen als auch aus ihren positiven Entwicklungen. Das Ergebnis wird in unserem Fall eine funktionsorientierte experimentelle Phonetik sein, die die verfügbaren Sprachverarbeitungstechniken im Forschungsfeld menschlicher Kommunikation einsetzt. Die Prosodieforschung hat hier eine ganz zentrale Aufgabe, und die Kieler Prosodiegruppe kann sich entsprechend profilieren.