

3.3 Schallsignalanalysen und Meßmethoden der Intonation in der akustischen Phonetik?

Um lautsprachliche Äußerungen zu erzeugen, müssen die Phonationsorgane (Stimmklappen/Stimmbänder und Stimmritze/Glottis) sowie die beweglichen Artikulationsorgane (Zunge, Lippen, Unterkiefer, Gaumensegel) zusammen den "exhalatorischen pulmonalen Luftstrom" (Kohler 1977: 58) modellieren. Zwei Positionen der Stimmklappen sind für die folgenden Betrachtungen besonders wichtig: (a) die Stimmklappen sind weit geöffnet (Position für die Produktion stimmloser Laute), (b) die Stimmklappen sind geschlossen und werden von der durchströmenden Luft in quasi-periodische Schwingungen versetzt (Produktion stimmhafter Laute). Je schneller die Stimmklappen schwingen, umso höher ist die Frequenz⁸ des artikulierten Lautes. Von entscheidender Bedeutung für die Intonationsanalyse ist die schon oft erwähnte Grundfrequenz (F_0), das akustische Korrelat des Intonationsverlaufs.⁹ Die Grundfrequenz ist nur bei Stimmhaftigkeit, also in den Phasen (quasi) periodisch schwingender Stimmklappen meßbar. Stimmlose Konsonanten weisen keine akustisch meßbare Grundfrequenz auf; das gleiche gilt auch für alle geflüsterten Laute.¹⁰ Als Grundfrequenz wird in der akustischen Phonetik die periodische Wiederholungsfrequenz einer komplexen Schallwelle bezeichnet. Anders gesagt:

"When a repetitive wave form is analysed into its component frequencies the fundamental (frequency) is the highest common factor of the component frequencies." (Ladefoged 1962: 111)

7 In diesem Abschnitt kann natürlich nur das für die folgenden Untersuchungen Wichtigste (und auch das nur gekürzt) dargestellt werden. Für detaillierte Einführungen in das Arbeitsgebiet der akustischen Phonetik sei auf die einschlägigen Kapitel in Kohler (1977), Ladefoged (1982) sowie Ladefoged (1962) hingewiesen.

8 Die Frequenz ist die akustische Eigenschaft eines periodischen Schallsignals, die sich aus der Anzahl der vollständigen Wiederholungen eines Schwingungszustands pro Zeiteinheit ergibt. Die Meßgröße ist Hertz (Hz). Die noch in einigen älteren Lehrbüchern (z.B. Ladefoged 1962) verwendete anglo-amerikanische Meßgröße cps (cycles per second) gibt diese Definition direkt wieder.

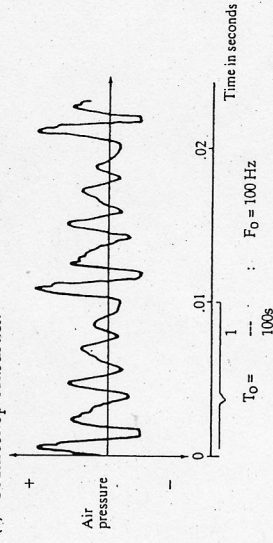
9 "Bei den für die menschliche Kommunikation relevanten Schallreizen handelt es sich nicht um reine Töne, sondern um komplexe Schallsignale, die, wenn sie periodisch sind, als 'Klänge' bezeichnet werden. Obwohl im Gegensatz zu reinen Tönen nicht immer ein eindeutiges akustisches Korrelat existiert, kann einem Klang perzeptiv eine bestimmte Tonhöhe zugeordnet werden." "Die Schwingungsform wird als annähernd periodisch und komplex bezeichnet, 'periodisch', weil sich das Muster der Schwankungen über die Zeit hinweg wiederholt, 'komplex', weil die Schwingungen nicht sinusförmig verlaufen." (Helfferich 1985: 76 und 38)

10 Geflüsterte Vokale in der Reihenfolge [l, i, ε, æ, a, u] bilden jedoch eine Serie von Lauten mit abfallender Tonhöhe. Wahrgenommen wird aber nicht die Grundfrequenz, sondern der zweite Formant (F_2) (vgl. Abb. (2) sowie Ladefoged 1982: 174)).

"Der zeitliche Abstand (...) zwischen einem positiven Nulldurchgang und dem nächsten positiven Nulldurchgang in Sekunden (s) ist die Grundperiode T_0 , deren Kehrwert (ausgedrückt in Hz) die Grundfrequenz F_0 ist." (Helfferich 1985: 38)

Die Abbildung (7) aus Ladefoged (1982: 166) zeigt die komplexe Schallwelle eines stimmhaften Lautes, des Vokals [ɔ]. Die durch die Schwingungen der Stimmklappen, die quasi-periodische Öffnung und Schließung der Glottis, verursachten Luftdruckveränderungen bei der Produktion dieses Lautes wurden mit einem Kathodenstrahl-Oszilloskop (Cathode Ray Oscilloscope) gemessen.

(7) Oszilloskop-Ausdruck



Der Laut aus Ladefogeds Beispiel wurde mit einer Grundfrequenz von 100 Hz realisiert, denn nach einer hundertstel Sekunde wurde der erste positive Nulldurchgang der komplexen Schwingung gemessen. Da die Aufzeichnung mit einem positiven Nulldurchgang beginnt, ist dieses Zeitintervall die Grundperiode T_0 .

Spektralanalysen mit einem Sonagrammen zeigen im Breitband-Sonagramm ebenfalls den Schwingungszyklus der Stimmbänder (senkrechte Linien). Darüber hinaus zeigt das Sonagramm aber auch die sich in bestimmten Frequenzbereichen verdichtende relative Intensität der einzelnen Schwingungskomponenten (waagerechte Schwärzungsbereiche), d.h., die für dieses Signal spezifische Platzierung der Formanten (Obertöne, Resonanzfrequenzen). Bei Kohler (1977) findet sich dazu Folgendes:

"Eine (...) Betrachtungsweise geht davon aus, daß das quasi-periodische Luftdrucksignal an der Glottis kein einfaches Signal ist, sondern in eine Reihe von einfachen Signalen zerlegt, analysiert werden kann, so wie ein Klavier'ton kein einfacher Ton ist, sondern mehrere einfache Töne enthält. Die einfachen Signale sind alle von derselben Form (Sinoidalsignale) und haben dieselbe Frequenz wie das komplexe Signal bzw. ganzzahlige Vielfache davon: Grundfrequenz und Oberschwingungen. Im Bereich der Wahrnehmung entsprechen

diesen Grundton und Obertöne (Harmonische). Das komplexe Signal wird so wie jedes streng periodische Signal nach dem Theorem von Fourier zerlegt (Fourieranalyse) und durch ein Spektrum dargestellt, das angibt, mit welcher Stärke (Amplitude) die einzelnen Frequenzen vertreten sind (Spektralanalyse)." (Kohler (1977: 51f))

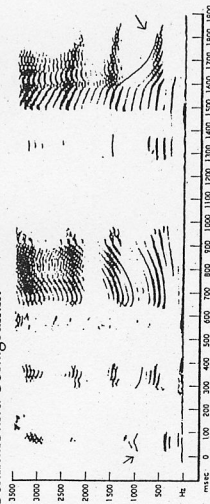
Besonders die drei untersten Formanten (F_1 , F_2 , F_3) sind in der Regel deutlich ausgeprägt im Sonogramm zu erkennen. Sie sind die Materialisierung unterschiedlicher Vokalqualität im akustischen Signal (vgl. auch Anm. 10) und korrespondieren mit dem unterschiedlichen Volumen der Resonanzräume, die im Ansatzrohr mit Hilfe der beweglichen Artikulationsorgane hergestellt werden.

Im Schmalband-Sonogramm werden die Frequenzen (Obertöne) einzeln sichtbar gemacht, aus denen sich die Formanten zusammensetzen. Da die Obertöne immer ganzzahlige Vielfache der Grundfrequenz sind, kann über den Verlauf z.B. des zehnten Obertons auf die Grundfrequenz geschlossen werden - sie entspricht einem Zehntel der Frequenz des zehnten Obertons. Die Abbildungen (8) und (9) (Ladefoged 1982: 192) zeigen von der Frage *Is Pat sad, or mad?* sowohl ein Breitband- als auch ein Schmalband-Sonogramm. Im Schmalband-Sonogramm ist der zehnte Oberton zusätzlich markiert.

(8) Breitband-Sonogramm



(9) Schmalband-Sonogramm

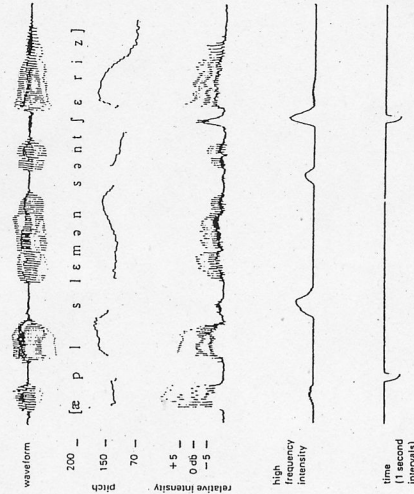


Beide Verfahren sind sehr zeitaufwendig und daher nur bei der Analyse von sehr kurzen Einheiten anwendbar.

Für längere Analyseeinheiten bietet es sich hingegen an, einen mit einem Tintenschreiber arbeitenden Oszillographen, einen Oszillomink, zu verwenden. Ein solches Gerät kann zwar nicht alle für das Schallsignal typischen schnellen

Luftdruckveränderungen aufzeichnen (der Tintenschreiber bewegt sich nicht schnell genug), aber es zeichnet Ausschläge auf, die jeder Vibration der Stimmbänder¹¹ entsprechen. Je mehr Ausschläge des Oszillomink pro Zeiteinheit, umso schneller schwingen die Stimmlippen und umso höher ist die Frequenz des artikulierten Lauts. Die Abbildung (10) eines Oszillomink-Ausdrucks mit seinen sechs Spuren für das Oszillogramm, eine auditiv-phonetische Transkription, den Grundfrequenzverlauf, zwei Intensitätskonturen und den Zeittakt ist ebenfalls Ladefoged (1982: 167) entnommen:

(10) Oszillomink-Ausdruck



Im Prinzip ist es also möglich, über das Zählen der Ausschläge die Frequenz des artikulierten Lauts festzustellen. In dem Beispiel (10) lassen sich im Oszillogramm ungefähr 12 Ausschläge bei der Produktion des Lautes [æ] zählen. Die Dauer des Lauts ist ungefähr eine Zehntel-Sekunde. Somit wurde dieser Laut mit einer Frequenz von ungefähr 120 Hertz produziert.

Zur Feststellung der Grundfrequenz wäre diese Meßmethode ebenfalls eine sehr zeitaufwendige. Für diesen Zweck werden deshalb sogenannte 'Tonhöhen-schreiber' (pitch meter) verwendet, die aber natürlich nicht die Tonhöhe, eine

11

Die Wellenform des Oszillominks (Abb. 10) zeigt deutlich die Übergänge zwischen stimmlosen (aperiodischen) und stimmhaften (quasi-periodischen) Lauten, auch eine weitergehende Segmentierung (Übergänge zwischen Nasalen /Liquiden und Vokalen bzw. stimmlosen u. stimmhaften Konsonanten) ist u.U. möglich. In stimmhaften Lauten entspricht die Anzahl der Ausschläge den Schwingungen der Stimmbänder.

auditive Wahrnehmungsgröße, sondern die Grundfrequenz messen (vgl. die dritte Spur des Oszillogramm-Ausdrucks). Im aufgezeichneten Grundfrequenzverlauf zeigt sich die Interaktion mindestens zweier, voneinander völlig unabhängiger Parameter: zum einen die für die folgenden Analysen zentrale phonologisch motivierte suprasegmentale Intonation (grammatisch relevante tonale Struktur), zum anderen die von der phonologisch-phonetischen Struktur abhängigen segmentalen Einflüsse¹² und Perturbationen¹³.

Für die folgenden Analysen ist es nützlich, verschiedene Ebenen der phonetischen Betrachtungsweise von Suprasegmentalia auch terminologisch voneinander zu trennen, um artikulatorische, akustische und auditive Eigenschaften und Prozesse nicht miteinander zu vermischen.

Die schon oft erwähnte **Tonhöhe** ist eine Eigenschaft der auditiven Wahrnehmung: Laute lassen sich in dieser Beschreibungsdimension auf einer relationalen Skala von tief bis hoch platzieren. Für eine artikulatorische Beschreibung heißt das, daß der subglottale Druck, der Öffnungsgrad der Glottis und die Spannung der Stimmlippen zusammenwirken müssen, um den Luftstrom in einer bestimmten Weise zu modellieren. Immer wenn ein Laut eine bestimmte wahrnehmbare Tonhöhe hat, wird die aus den Respirationsorganen ausströmende Luft durch die **Vibration der Stimmlippen** in regelmäßige (quasi-periodische) Schwingungen versetzt. Diese Schwingungen sollten idealiter keine Perturbationen durch die Artikulatoren erfahren. Kandidaten für tontragende Laute sind deshalb vor allem Vokale, Nasale und Liquide. Die Quote, mit der sich so ein quasi-periodisches Schwingungsmuster wiederholt, wird in der akustischen Phonetik die **Frequenz** des betreffenden Lautes genannt. Da es sich bei der Lautproduktion, wie in Abb. (7) gezeigt, stets um komplexe Schallwellen handelt, ist die der Tonhöhe entsprechende akustische Meßgröße die Grundfrequenz (F_0 , gemessen in Hz).

Ein zweiter wichtiger Parameter, mit dem Laute beschrieben werden können, ist die **Lautstärke**. Auch dies ist ein der auditiven Wahrnehmung zugeordneter Begriff: Laute lassen sich auf einer ebenfalls relationalen Skala von leise bis laut

12 Unter dem Stichwort "vowel intrinsic pitch" (intrinsische Tonhöhe von Vokalen) wurde u.a. von Lehiste (1970: 68) ein Zusammenhang zwischen Vokalqualität und durchschnittlicher Grundfrequenz festgestellt. Zungenhöhe Vokale weisen höhere durchschnittliche F_0 -Werte auf als zungen tiefe Vokale. Neuere Arbeiten dazu sind: Ladd & Silverman (1984), Steele (1986), Umeda (1981) und für das Deutsche Antoniadis & Strube (1981), Newekowsky (1975).

13 Hier ist vor allem der Einfluß der Konsonanten auf die Ausprägung der F_0 -Kontur zu nennen. Hypothese: Die F_0 -Kontur fällt nach stimmlosen Konsonanten und steigt nach stimmhaften Konsonanten, vgl. Lehiste & Peterson (1961) und Abramson & Lisker (1985). Kritische Analysen zu dieser Dichotomie, die vor allem auf den Einfluß der Intonation hinweisen, finden sich bei Beckman (1986), Silverman (1986) und Kohler (1982). Hirst (1983) weist jedoch darauf hin, daß eine von stimmlosen Segmenten unterbrochene Intonationskontur funktional äquivalent zu derselben Kontur ohne diese Unterbrechungen ist.

platzieren. Artikulatorisch entspricht dem Lautstärkeanstieg im wesentlichen ein Anstieg des Luftdrucks. Bezieht man sich auf den instrumentell meßbaren Faktor der Lautstärke, so spricht man von der Intensität eines Lautes. Die Intensität von Schallsignalen wird in Dezibel über die Amplitude der Schallwelle gemessen.¹⁴

Eine dritte auditiv wahrnehmbare Eigenschaft eines Lautes ist seine Länge. Artikulatorisch wird die Länge durch die Ausdehnung des Artikulationsprozesses erzielt. Auf der akustischen Beschreibungsebene entspricht dieser zeitlichen Dimension die Dauer eines bestimmten Schallereignisses.

(11) phonetische Analyseebenen

auditive Phonetik	artikulatorische Phonetik	akustische Phonetik
Tonhöhe Wahrnehmungs- skala: hoch - tief	quasi-periodische Schwingungen der Stimmblätter	Grundfrequenz (F_0) Meßgröße: Hertz (Hz)
Lautstärke Wahrnehmungs- skala: laut - leise	artikulatorischer Aufwand, Luftdruck	Intensität Meßgröße: Dezibel (db)
Länge Wahrnehmungs- skala: lang - kurz	Artikulations- prozeß	Zeit, Dauer Meßgröße: Millisekunde (ms)

3.4 Die akustische Realisierung des Akzents im Deutschen

Wenn man - wie in Kapitel 2 geschehen - systematisch zwischen einem abstrakten phonologischen Potential für Akzentuierbarkeit von Silben, genannt **Iktus**, und der tatsächlich realisierten Hervorhebung bestimmter Silben, genannt **Akzent**, unterscheidet, stellt sich natürlich die Frage, wie diese Hervorhebung in einer Sprache wie dem Deutschen phonetisch realisiert wird. Die Suche nach Zusammenhängen zwischen dem Eindruck von Akzentuiertheit und bestimmten akustischen Eigenschaften des Schallsignals hat eine lange Forschungstradition in der Phonetik.¹⁵ Doch (ähnlich wie die Suche nach der akustischen Manifestation der Silbe im Schallsignal) die Identifizierung der akustischen Korrelate des Akzents

14 Ramers (1988: 35) weist allerdings darauf hin, daß 'Dezibel' keine rein physikalische Meßgröße ist, sondern auf die Art der Wahrnehmung der Schallintensität durch das menschliche Ohr Bezug nimmt. Vgl. auch Lehiste (1970: 113ff).

15 Vgl. Lehiste (1970: Kap. 4).

erwies sich als problematisch.¹⁶

So finden sich bei Winkler (1984: 735) in der Dudengrammatik zwar einige Hinweise, die auf den ersten Blick auch eine akustische Grundlage zu haben scheinen, bei näherer Betrachtung jedoch rein impressionistisch sind:

"Unter schwer verstehen wir das Gewicht, mit dem eine Silbe in Ohr fällt. Die Rede bedient sich dabei verschiedener stimmlicher Mittel. Schwer wirkt eine Silbe vornehmlich, wenn sie entscheidende Stufen der Auf- oder Abschwünge des Tonfalls übernimmt (musikalischer Akzent). Daneben hebt meist zugleich vermehrte Lauthheit (Betonung durch Atemdruck) die Silbe heraus (dynamischer Akzent). Auch Dehnung des Vokals kann eine Silbe beschweren (temporaler Akzent). Anderes, z.B. eine Stauung des Redeflusses vor einem Wort, auch Klangwechsel, kann mißspielen. Gewöhnlich verbinden sich diese Beschwerungsweisen. Im Deutschen fällt, etwa im Vergleich mit dem Französischen, die Betonung durch Lauthheit auf; doch bleibt auch in der deutschen Rede der musikalische Akzent entscheidend."

Ähnlich äußern sich auch Neppert & Pétursson (1986: 145):

Akzentuierte Vokale haben meistens eine höhere Grundfrequenz als unakzentuierte. Eine derartige Erhöhung der Grundfrequenz als Kennzeichnung des Akzents ist sehr häufig. Sie wird auch im Deutschen verwendet, das häufig als eine Sprache mit Intensitätsakzent bezeichnet wurde. Im Deutschen, Französischen und Englischen sowie in den slawischen Sprachen ist die Erhöhung der Grundfrequenz auf der akzentuierten Silbe das weitaus wichtigste Element der Akzentuierung. In Sprachen wie Spanisch, Italienisch, Schwedisch, Norwegisch und Japanisch wird die Akzentuierung fast ausschließlich durch Erhöhung der Grundfrequenz bewirkt. Ein solches Akzentuieren wird häufig melodisches oder tonales Akzentuieren genannt."

Als physiologisches Korrelat perzipierter Akzente nennt Kohler (1977: 84f) zwar eine "größere Aktivität der gesamten Brustatmungsaktivität (...) als dies durch die Abnahme des Luftvolumens in der Lunge gefordert wird." Im akustischen Signal findet sich jedoch nach Kohler kein einfaches Korrelat, sondern Intensität, Dauer und Grundfrequenz wirken in sehr komplexer Weise zusammen.

"Offensichtlicher Kandidat für ein akustisches Korrelat der Betonung wäre die gemittelte Schallintensität des am Ausgang des Ansatzrohres abgestrahlten Signals, d.h. die Schallenergie, die im Mittel über eine festgelegte Zeitspanne hinweg durch die Flächeninheit strömt. Sie wird jedoch durch Phonation und supraglottale Artikulationen sowie durch die Höhe des mittleren subglottalen

¹⁶ Ähnlich äußert sich auch Klein (1980: 32, Anm. 22): "Es scheint mir kein gutes Zeichen für den Stand der deutschen Phonetik, daß völlige Unklarheit darüber herrscht, was das signalphonetische Gegenstück des 'Akzents' ist."

Drucks (als mittelbarer Auslöser einer bestimmten Lautstärke beim Hörer) ebenso bestimmt wie durch einen zusätzlichen subglottalen Druckanstieg (als Anzeige von Akzentuierung). Eine Trennung dieser verschiedenen Komponenten allein in der Aufzeichnung der Schallintensität ist schon deshalb schwer möglich, weil zunächst noch gar nicht bekannt ist, welche Mischungsverhältnisse optimal wären. Außerdem wird der Druckanstieg zum Zwecke der Akzentuierung auch in eine höhere Frequenz der Stimmbandschwingung sowie in andere Dauerverhältnisse artikulatorischer Segmente - neben einer größeren Schallintensität - abgebildet, so daß die intendierte Akzentuierung im akustischen Signal nicht durch einen physikalischen Parameter repräsentiert wird und die Betonungswahrnehmung sich eben auch auf diese drei akustischen Größen stützt." Kohler (1977: 85)

Dies ist zwar vermutlich richtig und angemessener als die Darstellungen von Winkler und Neppert & Pétursson, doch scheinen die verschiedenen akustischen Parameter bei der Akzentuierung nicht nur in der von Kohler geschilderten, recht komplizierten Weise zusammenzuwirken, sondern auch kompensatorische Funktionen übernehmen zu können und darüberhinaus unterschiedlich gewichtet werden zu können.

Eine solche Gewichtung findet sich bei Isačenko & Schädlich (1966). Die Autoren betrachten Intensität und Grundfrequenz als selbständige und unabhängige Größen, wobei sie "(a) die Wirkung von Intensitätsunterschieden bei konstanter Grundfrequenz einerseits und (b) die Wirkung von Grundfrequenzunterschieden bei (...) konstanter Intensität andererseits" untersuchen. Für die experimentelle Untersuchung wurden zunächst einfache deutsche Sätze gesprochen, die anschließend auf zwei Frequenzen monotonisiert wurden, damit ein 'hochtoniges' und ein 'tief-toniges' Band zur Verfügung standen.¹⁷ Für die Perzeptionsexperimente wurden Segmente aus beiden Bändern zusammengeschnitten. Isačenko & Schädlich (1966: 14f) weisen darauf hin, "daß die im Signal vorliegenden Tonbrüche nicht als solche wahrgenommen werden: das Ohr nimmt den Tonbruch als natürliche Transition auf." Zur Gewichtung der beiden Parameter werden die Wörter *Übersetzen* und *überSETzen* mit den unterschiedlichen Akzentpositionen produziert und anschließend jeweils hochtonig und tief-tonig monotonisiert. In den monotonisierten Formen sind die Akzentpositionen eindeutig erkennbar. Die akustische Hervorhebung ordnen die Autoren dem Parameter Intensität zu. Bei Umkehrung der ursprünglichen Grundfrequenzverhältnisse wurden 50 Versuchs-

¹⁷

Das tief-tonige Band wurde bei 150 Hz und das hochtonige bei 160 bzw. 178, 6 Hz monotonisiert. Der Abstand von ca. 10 Hz entspricht bei diesem Frequenzausschnitt einem Halbtonschritt, der nach Ansicht der Autoren zugleich das Minimum und auch das Optimum für eine sichere perzeptive Unterscheidung der beiden Tonstufen darstellt. Riefeld & Gussenhoven (1985) kommen für das Holländische auf etwas größere Intervalle; sie weisen auch darauf hin, daß die Hertzskala der Prominenzwahrnehmung besser entspricht als die Halbtonskala.

personen die folgenden Stimuli zur Identifizierung der Akzentsilbe vorgelegt:

(12a) ü | berSEtzen

(12b) ÜBer | set | zen

85% identifizierten die Silbe *ü-* im Beispiel (9a) und 98% die Silbe *-set-* als hervorgehoben. Aus diesem Test schließen die Autoren, daß im Deutschen Intensitätsveränderungen für die Perception von Hervorhebung sekundär sind. Sie können diese Funktion zwar kompensatorisch in Abwesenheit von Grundfrequenzveränderungen zwar übernehmen, wenn jedoch Intensitäts- und F_0 -veränderung vorhanden sind, aber nicht dieselbe Silbe markieren, ist der Tonbruch der entscheidende, primäre Parameter.

In die gleiche Richtung zielen auch die Ergebnisse der Experimente, die Fry (1955, 1958) zum Englischen auf der Basis von komplexen Perceptionstests mit synthetisierten Stimuli durchgeführt hat, die es ihm ermöglichten, die Parameter *Dauer*, *Intensität* und *Grundfrequenz* separat und kontrolliert zu variieren. Der Einfluß der F_0 -Veränderung wurde in zwei Formen untersucht: Zum einen wurden, vergleichbar mit Isaenko und Schädlichs Tests, statische Tonbrüche zwischen zwei Silben betrachtet; zum anderen komplexere Muster von F_0 -Veränderungen mit steigenden und fallenden Konturen innerhalb von Silben.¹⁸ Testbasis waren Paare von Nomen und Verb des Typs *Object* vs. *object*, die den Versuchspersonen im Kontext von Trägersätzen (*Where is the accent in —*) präsentiert wurden. Mithilfe von differenzierten Auswertungsmethoden ordnet Fry die drei untersuchten Parameter in absteigender Relevanz für die Perception von Akzentuiertheit¹⁹ folgendermaßen an:

- 1.) F_0 -Veränderungen
- 2.) Dauer
- 3.) Intensität

Sowohl *Dauer*- als auch Intensitätsveränderungen beeinflussen die Akzentwahrnehmung: Wenn ein Vokal im Vergleich zum anderen längere Dauer oder höhere Intensität aufwies, wurde dieser von der Mehrzahl der Versuchspersonen als akzentuiert wahrgenommen; der Effekt verstärkte sich noch, wenn Dauer und Intensität parallel verändert wurden. Bei konstanter Intensität und steigender Dauer im Gegensatz zu konstanter Dauer und steigender Intensität erwies sich der

¹⁸ Diese komplexen Muster bezeichnet Fry als "sentence intonation".

¹⁹ Fry unterscheidet nicht zwischen "stress" und "accent"; in seinen Untersuchungen findet sich das hier als Akzent bezeichnete Phänomen unter dem Stichwort "stress".

Parameter *Dauer* als der für die Akzentwahrnehmung im Englischen wichtigere. Die Experimente mit den Tonbrüchen zeigen, daß eine hochtonige Silbe eher als akzentuiert wahrgenommen wird als eine tieftonige und daß F_0 -Veränderungen innerhalb von Silben den stärksten Einfluß auf die Akzentwahrnehmung haben. Frys Analysen zufolge unterscheiden sich jedoch F_0 -Veränderungen jeder Art von Veränderungen der Parameter *Dauer* und *Intensität* dergestalt, daß sie einen 'Alles-oder-Nichts'-Effekt produzieren, indem die Größe der F_0 -Veränderung, so sie die Schwelle der Perzipierbarkeit überschritten hat, weniger wichtig zu sein scheint als die Tatsache, daß überhaupt eine Veränderung stattgefunden hat.

Ein Problem sowohl bei Isaenko & Schädlich als auch bei Fry ist jedoch die Künstlichkeit des untersuchten Materials. Hier manifestiert sich ein generelles Dilemma der Experimentalphonetik. Es läßt sich nämlich aus der Eliminierung von bestimmten Parametern und der systematischen Variation anderer Eigenschaften des Signals nicht folgern, daß die eliminierten Parameter perzeptiv irrelevant sind. Vielmehr ist damit zu rechnen, daß der kompetente Hörer Fehlendes aus anderen signalphonetischen Ressourcen oder aufgrund seines sprachlichen oder kontextuellen Wissens rekonstruiert. So weist Kohler (1977: 83) darauf hin, daß

"das menschliche Gehör auch in der Lage ist, bei Fehlen der Grundfrequenz (wie z. B. in der Telefonübertragung) über die Periodizität des Signals zur selben Tonhöhenwahrnehmung zu kommen, wie wenn die Grundfrequenz vorhanden wäre."

Aus diesem Befund sollte allerdings nicht geschlossen werden, daß die Grundfrequenz ein irrelevanter Parameter für die Perception der Tonhöhenwahrnehmung ist. Ähnlich Probleme sind bei der Suche nach den relevanten akustischen Korrelaten der perzipierten Akzentuierung zu erwarten.

Eine andere Möglichkeit, sich dem Untersuchungsgegenstand zu nähern, geht daher nicht von der Perception synthetischer Stimuli, sondern von der Produktion natürlicher Daten aus und versucht z.B., die akustischen Eigenschaften von akzentuierten Silben über den Vergleich mit unakzentuierten Silben zu bestimmen. Ein solches Experiment zum Japanischen und Englischen wird bei Beckman (1986: Kap. 6) vorgestellt, das auch den gestiegenen Ansprüchen an die technische Aufbereitung der Versuchsanordnung gerecht wird. Sie untersucht Wörter mit verschiedenen Akzentmustern (für das Englische werden die gleichen Minimalpaare wie in Frys Experimenten benutzt) und versucht eine Gewichtung der akustischen Parameter nicht über die Perception synthetisierter Stimuli, sondern über die Mög-