

NASALE UND ORALE VOKALE — STRUKTUR UND PERZEPTION

GOTTFRIED MEINHOLD*

Nach ihrer spektralen Struktur sind die nasalen Vokale gegenüber den oralen vor allem durch eine Schwächung der Formanten F_1 , F_2 und F_3 (Smith, zit. bei Fant), durch das Auftreten von zusätzlichen Intensitätskonzentrationen oberhalb 1 kHz, durch einen Antiformanten bei 500 Hz (Hattori, Yamamoto und Fujimura) sowie durch eine relative Zunahme der Grundtonamplitude charakterisiert. Wenig beachtet wurde in der Literatur bisher der geringere Betrag der Gesamtintensität; Nasalvokale zeigen verminderte spezifische Schallfülle. Die genannten spektralen Erscheinungen — nämlich Abbau vokaldifferenzierender Teiltongruppen, Auftreten neuer oder vermehrter gemeinsamer Intensitäten vor allem im Bereich zwischen den Formanten — bewirken eine Nivellierung der Formantstruktur, eine Abtragung des scharf umrissenen Formantprofils. Der Grad der Konzentration der Intensität auf bestimmte Frequenzbereiche nimmt ab, die Intensitätsdispersion erhöht sich. Im Vergleich zu den entsprechenden oralen Vokalen büßen die nasalierten dadurch erheblich an spezifischer Klangcharakteristik ein und damit an Fähigkeit, untereinander gut wirksame Kontraste zu bilden; sie verlieren an Gestalt, weil der Grad ihrer Strukturiertheit abnimmt, ihre Identifizierung wird erschwert.

Normale Vokale ($[a, e, i, o, u]$) und ihnen entsprechende, im Sinne einer Rhinolalia aperta nasalierte, und zwar mit einem Grundton von ca. 113 Hz (männl. Vp), versuchte Vf. zunächst in ihrer Intensitätsdispersion mit Hilfe der Shannonschen Entropie zu vergleichen. Die Negentropie dient unter bestimmten Voraussetzungen als allgemeines statistisches Streuungsmaß. Demzufolge berechnet man das spektrometrische Bild mit den Frequenzklassen auf der Abszisse und der Intensität auf der Ordinate als relative Häufigkeitsverteilung und erhält für den Kompaktheitsgrad (Halle und Jakobson) eine Maßzahl, die als Vergleichswert für Vokale dienen kann, welche mit gleichen Spektrometern und gleichen Stimmen bei gleichmäßiger Aussteuerung und gleicher Grundfrequenz gewonnen wurden. Die spektrale Entropie dient also nicht als absoluter klanganalytischer Parameter, sondern als Vergleichsmaß in relativ engen Grenzen. Die Zunahme der Intensitätsstreuung bei nasalen Vokalen kann man gut an den Quotienten ablesen, die aus beiden Werten zu bilden sind

* Abt. Sprechwissenschaft der Universität, Jena (DDR).

$\left(\frac{H_{nas}}{H_{or}}\right)$. Der Vokal *a* wurde durch die nasopharyngeale Öffnung akustisch nur gering beeinflusst, zeigt also eine geringe Zunahme der spektralen Entropie, bei *i* und *o* dagegen ist sie erheblich.

Tabelle 1

	<i>a</i>	<i>e</i>	<i>i</i>	<i>o</i>	<i>u</i>
H_{oral}	4,2138	3,8923	3,4123	3,4923	4,0675
H_{nasal}	4,2389	4,3249	4,0356	4,1598	4,5681
$\frac{H_{nas}}{H_{or}}$	1,01	1,11	1,18	1,19	1,12

Der Entropiequotient blieb für *e*, *i*, *o*, *u* auch bei Vokalspektren von Hattori, Yamamoto und Fujimura, die Vf. miteinander verglich, in den Grenzen >1,09 und <1,19. — Die Berechnung der spektralen Entropie könnte u. a. als Objektivierungsmethode für Behandlungserfolge bei Rhinolalia bzw. Palatolalie (-phonie) dienen.

Im Perzeptionsversuch wurden zunächst Segmente derselben Vokale zu je 0,05 bis 0,06 sec Dauer von phonetisch vorgebildeten Hörern dekodiert (somit standen den Hörern über die Vokale ihrer Muttersprache hinaus Urteilsklassen zur Verfügung). Die Kürze sollte die Hörunsicherheit steigern. Die Registrierung der Imitationsprodukte bzw. der Urteile erfolgte durch den Vl.

Von oralen Vokalen wurden [*a e i*] sicher gehört, [*o*] dagegen teils als ,*u*' mit der geringen Urteilsunsicherheit von $H = 0,4999$ bt, [*u*] andererseits teils als ,*o*' und ,*m*' ($H = 0,5920$ bt).

Bei den nasalierten Vokalen bleibt [*ã*] sicher. In der Konfusionsmatrix (Tab. 2) werden daher nur die Urteile von [*i ã õ ü*] aufgeführt. Die aus den relativen Häufigkeiten der Konfusionsmatrix berechneten Urteilsentropien lauten (bei einer maximal zehn Urteilsklassen entsprechenden Entropie von $H_{max} = 3,3220$ bt):

	<i>i</i>	<i>õ</i>	<i>ü</i>	<i>ẽ</i>
H	2,3627	1,9152	1,8021	1,4464

(in bit)

Allerdings sagen diese Zahlen nichts über die qualitative Verschiebung bei der Dekodierung aus. [*i*] wird nämlich dann mit dem geringsten Verfehlungsindex (Tarnóczy) gehört, der durch die Summe der relativen Häufigkeiten aller anderslautenden Urteilsklassen ausgedrückt wird. (Da [*i*] zu 32,2 % i-ähnlich gehört wird, beträgt der Verfehlungsindex $100 - 32,2 = 67,8$ %.) Die aus dem Verfehlungsindex sich ergebende Rangfolge lautet also *i*, *ü*, *õ*, *ẽ* (vgl. Tab. 2).

Tabelle 2 Konfusionsmatrix (Angaben in %)

Urteile	, <i>e</i> '	, <i>i</i> '	, <i>o</i> '	, <i>u</i> '	, <i>y</i> '	, <i>æ</i> '	, <i>ø</i> '	, <i>m</i> '	, <i>a</i> '	, <i>ε</i> '
<i>ẽ</i>	6,7					29,4	58,9	5,0		
<i>ĩ</i>	13,0	32,2		25,2			7,8	15,7	6,1	
<i>õ</i>			13,3	15,6		12,2	34,4		11,1	13,3
<i>ü</i>				17,2	6,9	5,8	56,3	13,8		

Eine weitere Möglichkeit, den Kontrastverlust der abgehörten Nasalvokale untereinander zu messen, ergibt sich, wenn man die Vokale paarweise miteinander konfrontiert und die relativen Häufigkeiten der Urteile, die das jeweilige Vokalpaar gemeinsam hat, summiert (vgl. dazu auch Tab. 2, der die Zahlen entnommen sind):

Tabelle 3

Stim.	Gemeinsame Urteile	In %	Stim.	Gemeinsame Urteile	In %
<i>ẽ</i>	<i>e/a/m</i>	70,6	<i>ĩ</i>	<i>a/u/a</i>	39,1
<i>ĩ</i>	<i>e/a/m</i>	36,5	<i>õ</i>	<i>a/u/a</i>	84,6
<i>ẽ</i>	<i>a/æ</i>	88,3	<i>ĩ</i>	<i>u/m/a</i>	48,7
<i>õ</i>	<i>a/æ</i>	46,6	<i>ü</i>	<i>u/m/a</i>	87,3
<i>ẽ</i>	<i>a/m/æ</i>	93,3	<i>õ</i>	<i>a/u</i>	51,0
<i>ü</i>	<i>a/m/æ</i>	75,9	<i>ü</i>	<i>a/u</i>	73,5

Aus den Mittelwerten der drei für jeden Nasalvokal ermittelten Prozentzahlen ergibt sich folgende Rangfolge der Kontrastaufhebung: [*i*] 41,4 %; [*õ*] 60,7 %; [*ü*] 78,9 %; [*ẽ*] 84,1 %.

Derselbe Versuch, mit Vokalsegmenten doppelter Dauer (0,12 sec) wiederholt, ergab eine nur gering verbesserte Identifizierungssicherheit. Die Versuche werden mit längeren Segmenten fortgesetzt. Der eklatante Kontrastabbau der Nasalvokale bei Kürzung könnte als perzeptives Argument neben anderen (lautphysiologischen) Argumenten zur Erklärung von Tendenzen zur Denasalisierung in rezenten Sprachen (Hindi, Frz., Poln.) dienen. Hinsichtlich der kommunikativen Leistungsfähigkeit — soweit sie auf der Kontrastfähigkeit der Lautsignale beruht — kann man eine Rangfolge von den optimalen geschlossenen Vokalen über die sog. offenen bis zu den perzeptiv ungünstigsten nasalierten Vokalen aufstellen. Man muß also die perzeptiven Grenzen der Phonemunterscheidung ergänzend zu *Winkels* Untersuchung (1964) in engster Abhängigkeit von der spektralen Struktur der jeweiligen Vokale bzw. Laute definieren.

- Fant, G.: Acoustic Theory of Speech Production. s'Gravenhage 1960.
- Halle, M. u. Jakobson R.: Grundlagen der Sprache. Berlin 1960.
- Hattori, Sh., Yamamoto K. und Fujimura, O.: Nasalization of Vowels and Nasals. *Bull. Kobayasi Inst. Physic. Res.* 6 (1956) 226—235.
- Tarnóczy, A.: Verständlichkeits-Stabilität, Konfusionsmatrix und Verfehlungsindex. *ZPSK* 17 (1964) 333—338.
- Winkel, F.: Perzeptive Grenzen der Phonemunterscheidung. *Proceed. 5th Intern. Congr. Phonet. Sc.*, Münster 1964. Basel/New York 1965, 582—588.

DISCUSSION

Mangold:

1. Warum wurden die Vokale [I Y U Ō] nicht mitverglichen?
2. Vom Gesichtspunkt der Sprachuniversalien ist Entnasalierung nicht häufiger als Nasalierung.

Meinhold:

Warum nicht andere Vokale außer *a, e, i, o, u* untersucht wurden, hat vor allem arbeitstechnische Gründe. Es galt, zunächst an Hand einiger Kardinalvokale durch möglichst viele Entropieberechnungen die Verwendbarkeit des Entropiemaßes prinzipiell zu testen. Umlaute tauchen lediglich als Urteilklassen auf. Sicherlich könnte die erschwerte Unterscheidbarkeit der Nasalvokale auch ein Grund für ihre geringe Häufigkeit in den Sprachen der Welt sein, auch wenn hierfür ebenfalls lautphysiologische Gründe geltend gemacht werden können. — Die Versuche sollen bis zur Auffindung neuer spezifischer perzeptiver Grenzen der Phonemunterscheidung fortgesetzt werden, also neben geschlossenen und nasalen Vokalen auch mit sog. offenen.