

ÜBER DAS INDIVIDUELLE SPRACHSPEKTRUM

THOMAS TARNÓCZY

Mit einer neuen Methode ist es uns in 1956 gelungen,¹ leicht und rasch ein durchschnittliches Spektrum mehrerer Sprecher zu bestimmen. Das Wesentliche an dieser Methode ist, daß mehrere Personen zugleich in derselben Sprache verschiedene Texte sprechen und die Auswertung erfolgt auf Grund der superponierten Pegeldiagramme (Abb. 1). Diese Methode wurde in der Literatur von Meyer-Eppler als "Sprechchor-Methode" genannt.²

Mit diesem Kunstgriff kann ein zeitlich ständiges Geräusch mit kontinuierlichem Spektrum erzeugt werden, dessen Analyse mittels Bandfilter leicht möglich ist. Das so erreichte Spektrum widerspiegelt die spektrale Eigenart des Sprechens genauer, als eine Statistik von Spektren des langdauernden geläufigen Sprechens mehrerer Personen nacheinander.

Ein neues Untersuchungsproblem war die Bestimmung des individuellen Sprachspektrums. Der Grundgedanke der Lösung ist Folgender³: Was vorher mit mehreren Personen in gleicher Zeit geschehen ist, das wurde in der Zeit nacheinander mit einer einzigen Person verrichtet. Das heißt, man nahm das Sprechen einer Person mehrmals nacheinander auf und spielte dann die einzelnen Aufnahmen gleichzeitig wieder. Dann konnte dasselbe Sprachgeräusch gehört werden, wie im Falle viele Personen gleichzeitig sprechen. Zur praktischen Lösung wurden neun, 2 Minuten lange Textabschnitte gewählt. Jeder Text wurde auf ein einzelnes Magnetband aufgenommen, wonach die neun Bänder auf ein gemeinsames Band zusammenkopiert wurden. Das aufgenommene Sprachgeräusch wurde auf die übliche Weise analysiert.

Die Versuche zum Studieren des individuellen Sprachspektrums wurden von 3 männlichen und 3 weiblichen Personen verrichtet. Die Wahl erfolgte auf Grund irgendeiner, mit dem Ohr wahrnehmbarer Spracheigenart. Das Wesentliche des Untersuches war, dass man einen Zusammenhang zwischen den, mit dem Ohr wahrnehmbaren Eigenarten und dem objektiven Sprachspektrum findet. Die in-

¹ Tarnóczy, Th., *Folia Phoniatica*, 8, 65-70 (1956); *J. Acoust. Soc. Am*, 28, 1270-1275 (1956); *Acta Phys. Hung.*, 7, 87-105 (1957); *Acustica*, 8, 392-395 (1958).

² Meyer-Eppler, W., *Grundlagen und Anwendungen der Informations-theorie* (Springer, Berlin, 1959), S. 398.

³ Tarnóczy, Th., *Gedenkbuch von D. Pais* (Budapest, 1956), S. 117-125 (ungarisch).

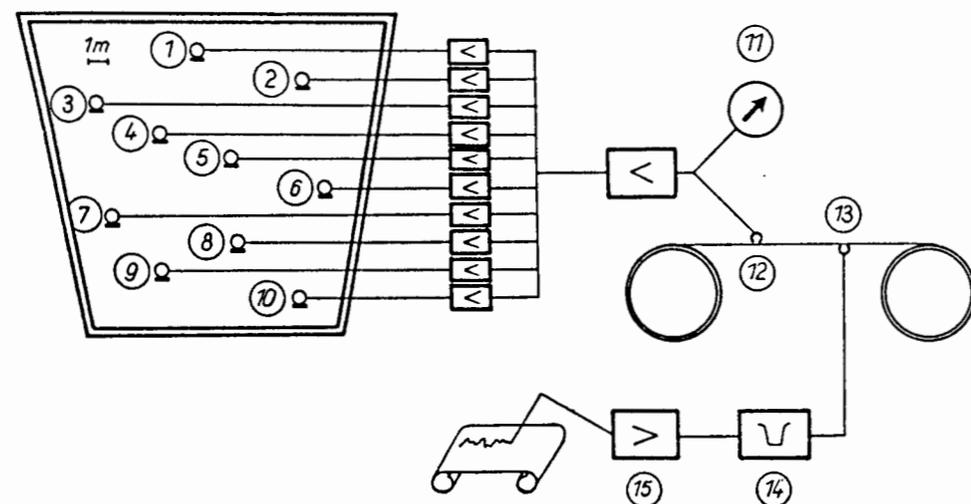


Abb. 1. Versuchseinrichtung zur Untersuchung der Sprachspektren. 1-10: Mikrophone; 11: Aussteuerungsmesser; 12: Aufnahmekopf; 13: Abspielkopf der Magnettonbandes; 14: Oktavfilter; 15: Pegelschreiber.

dividuellen Sprachspektren der Versuchspersonen wurden mit dem durchschnittlichen ungarischen männlichen bzw. weiblichen Sprachspektrum normaler Lautstärke verglichen. Die Analyse der einzelnen Spektren zeigt folgende Ergebnisse:

1. *männliche Stimme* (Abb. 2). Die Charakteristik der Männerstimme, d.h. das Maximum zwischen 80-700 Hz, kommt hier ausdrucksvoller als im durchschnittlichen Spektrum zur Geltung. Dieser Teil ist auffallend angestiegen und zeigt in Mittellage von 250 Hz einen bemerkbaren (2,5 dB) Abfall, wodurch ein zweifaches Maximum entsteht. Wahrscheinlich entspricht diese akustische Eigenart dem subjektiven klangvollen Charakter. Es ist interessant, daß auch der Teil des Spektrums über 1000 Hz höher liegt als im durchschnittlichen. Das zeigt sich am ehesten um 3000 Hz und ist wahrscheinlich mit der stärkeren Aussprache einzelner Engelaute verbunden. Um 3000 Hz hat der ungarische Engelaute [j] ein Maximum.

2. *männliche Stimme* (Abb. 3). Das bei dem vorhergehenden Falle beschriebene flache Maximum zeigt sich hier auch, sogar etwas tiefer verschoben: 70-700 Hz. Das Maximum ist nicht so hervorspringend und nicht zweifach, es steht also dem durchschnittlichen Spektrum näher. Zwischen 500-2500 Hz läuft das Spektrum der 2. männlichen Stimme mit dem durchschnittlichen weiblichen Spektren zusammen. Es ist höchstwahrscheinlich, daß diese Eigenschaft mit dem weicheren Charakter der Stimme in Verbindung ist. Das Spektrum im Bereich zwischen 500-2500 Hz hat einen Abfall von 18 dB, im Gegensatz zum durchschnittlichen von 22, 5 dB und zur 1. männlichen Stimme von 22 dB.

Weiteres physikalisches Äquivalent der weichen Stimme ist die Schwächung der Bestandteile über 5000 Hz. Hier fällt das Spektrum der 2. männlichen Stimme unter



Abb. 2. Das Sprachspektrum der 1. Männersprache (ausgezogene Kurve) und das durchschnittliche normale männliche Sprachspektrum (gestrichelte Kurve) in relativen dB/Hz-Werten.



Abb. 3. Das Sprachspektrum der 2. Männersprache (wie in Abb. 2).

das durchschnittliche, d.h. es enthält Engelaute nur in kleinerer Intensität. Das bezieht sich besonders auf [s], des eben hier sein Maximum hat.

3. *männliche Stimme* (Abb. 4). Ein von dem bisherigen ganz verschiedenes Spektrum. Ihre erste auffallende Eigenschaft ist der nach dem scharfen Maximum entstehende rasche Abfall. Wenn dieses Spektrum mit dem durchschnittlichen weib-



Abb. 4. Das Sprachspektrum der 3. Männersprache (wie in Abb. 2).

lichen verglichen wird, so erkennt man einen gewissen "weiblichen" Charakter in der Kurve. Die Stimme ist zwar fast mit einer Oktave tiefer als die weibliche, der Abfall beginnt aber bei 130 Hz, im Gegensatz zu der vorher geschilderten Lage, wobei er nach 400 Hz begonnen hat.

Das zweite Maximum ist in der Gegend von 350 Hz eben zu erkennen. Der Abfall zwischen 500–2500 Hz ist nun auch 18 dB, welchen wir schon als Ausdruck der Weichheit der Sprachstimme erkannt haben. Über 3000 Hz fällt die Kurve unter das durchschnittliche männliche Spektrum, es zeigt also ein scharfes Fehlen der Engelaute. Die Versuchsperson hat anstatt des [s] und [z] Lautes einen, dem englischen [θ] und [ð] ähnlichen Laut gesprochen. Sowohl der "Sack" zwischen 4000–5000 Hz, wie auch der über 3000 Hz eintretende grosse Abfall ist mit dieser Tatsache im Zusammenhang.

1. weibliche Stimme (Abb. 5). Das Spektrum hat tiefere Komponenten als beim durchschnittlichen. Dies und das Maximum um 200 Hz zeigt an, dass man einer tief liegenden weiblichen Stimme gegenübersteht. Diese Eigenart wird noch dadurch gesteigert, daß die Kurve bei der hervorgehobenen Stelle des durchschnittlichen weiblichen Spektrums von 700 Hz tiefer als das durchschnittliche vorläuft. Das bedeutet auch die gleichzeitige Verminderung einiger Vokalformanten.

Auf diesem Gebiet ist das [a] und [e] von Interesse und tatsächlich – nachträglicher Kontrolle gemäß – hat sich der Laut [a] der Versuchsperson gegen [ä] verschoben, der [e] Laut war dagegen viel geschlossener als gewöhnlich. Also ist der untere Formant des [a] etwas tiefer, der des [e] etwas höher als im allgemeinen. Dieses Beispiel des sprachlichen Benützens der Methode halten wir als besonders wichtig, weil es sich um das Erkennen solcher Eigenarten auf Grund des Sprachspektrums



Abb. 5. Das Sprachspektrum der 1. Frauensprache (ausgezogene Kurve) und das durchschnittliche normale weibliche Sprachspektrum (gestrichelte Kurve) in relativen dB/Hz-Werten.

handelt, die auf Grund oberflächlichen Hörens nicht auffällig waren. Das Übereinstimmen bei der nachträglichen Kontrolle, hat selbst die Forscher überrascht. Beginnend von 100 Hz hat die 1. weibliche Stimme einen mit den durchschnittlichen ganz gleichen spektralen Ablauf. Ihr Abfall zwischen 500–2500 Hz beträgt 19 dB, mit dem durchschnittlichen ebenfalls übereinstimmend.

2. weibliche Stimme (Abb. 6). Diese Stimme hat einen mittelhohen Charakter, ihr Maximum fällt in die Gegend von 225 Hz. Das früher erwähnte Fehlen des Vokalformanten zeigt sich hier im grösseren Masse. Das Spektrum dieser Stimme bleibt zwischen 400–1000 Hz unter den durchschnittlichen. Dies kann natürlich auch mit dem Ohr gut wahrnehmbar sein: das Vermindern der Eigenschaft von den Vokalen [a], [ä] und [e] ergibt das Urteil der "dünnen" Klangfarbe. Genauere Kontrolle erfolgte nicht. Eine viel wichtigere Eigenart dieses Spektrums ist das Steigen bei 2800 Hz. Der Verfasser hat in früheren Untersuchungen⁴ die charakteristische Stelle der Nasalität in der Umgebung von 2500 Hz bezeichnet. Diese Stimme, deren nasale Klangfarbe zweifellos zu bestimmen war, zeigte gleichfalls um diese Stelle spektrale Unregelmässigkeiten.

3. weibliche Stimme (Abb. 7). Von den drei letzten Stimmen hat diese die härteste Klangfarbe, ihr Maximum fällt auf 250 Hz. Weitere wichtige Eigenart des Spektrums ist der Abfall über 3000 Hz. Der Ausprung bei 2000 Hz ist eine relative Folge der sich vor- und nachher zeigenden Abfälle, welche wieder mit dem schwächeren Aussprechen der Engelaute im Zusammenhang sind.

⁴ Tarnóczy, Th., *Word*, 4, 71–77 (1948).



Abb. 6. Das Sprachspektrum der 2. Frauensprache (wie in Abb. 5).

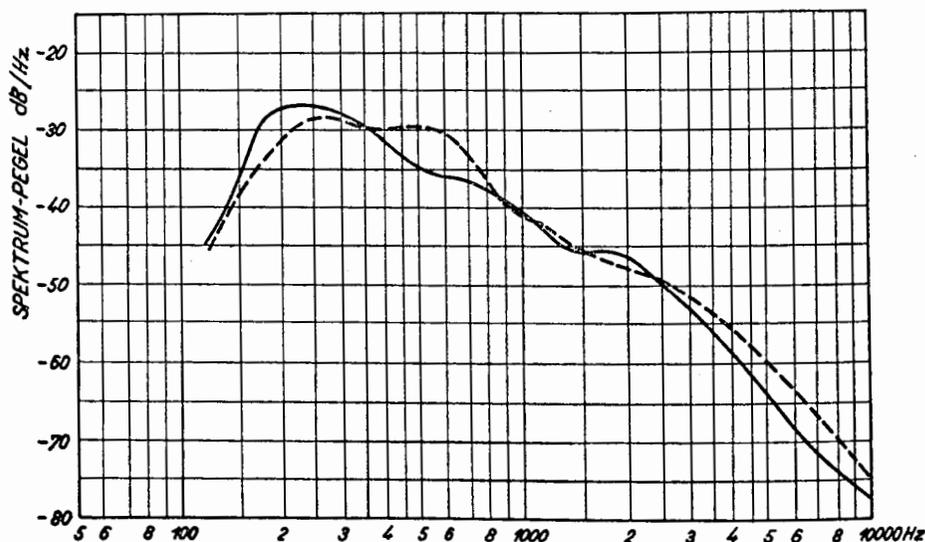


Abb. 7. Das Sprachspektrum der 3. Frauensprache (wie in Abb. 5).

Die obenstehenden Analysen sind Beispiele, mit welchen wir die Brauchbarkeit der Methode beweisen wollten. Man behauptet nicht, daß mit der Sprechchor-Methode alle Fragen der Sprechanalyse vollkommen gelöst angesehen werden können, man hofft aber wesentlich näher dazu gekommen zu sein.

*Ungarische Akademie der Wissenschaften
Forschungsgruppe für Akustik
Budapest*