

# Gibt es lange und kurze Vokale?

.....  
Dr. Hans-Otto Reuss und Fred Warnke

Die meisten lerntheoretischen Überlegungen hinsichtlich des Spracherwerbs gehen davon aus, dass zum Bestimmen kurzer oder langer Vokale eine Art Zeitgefühl notwendig ist. Auf der anderen Seite ist der in dieser Richtung arbeitenden Wissenschaft bekannt, dass die im Schulunterricht und auch im LRS-Förderunterricht benutzten Bezeichnungen so genannter „kurzer“ und „langer“ Vokale genau genommen nicht vertretbar sind. Nicht die zeitliche Länge ist für die Wahrnehmung entscheidend, sondern die Frequenzstruktur. Weil diese Tatsache aber weder in Kreisen des lerntheoretisch begründeten Rechtschreibaufbaues noch bei vielen Sprachheilpädagogen und Logopäden bekannt zu sein scheint, ist der nachstehende Fachartikel als ehrliches Bemühen zu verstehen, dieses Wissen mittels eines einfachen Beispiels auch in denjenigen Kreisen zu verbreiten, die es praktisch benutzen (müssten). Die Verfasser möchten an einem einfach nachvollziehbaren Experiment zeigen, dass es kein zeitlicher Parameter ist, der uns Menschen befähigt, „lange“ oder „kurze“ Vokale zu unterscheiden, sondern die Fähigkeit, unterschiedliche Frequenzstrukturen der Vokale wahrzunehmen. Hieraus wird gefolgert, dass bei Kindern, die nicht in der Lage sind, „lange“ oder „kurze“ Vokale voneinander zu unterscheiden, mit ziemlicher Wahrscheinlichkeit das Tonhöhenunterscheidungsvermögen nicht altersgerecht entwickelt ist.

Mit der zunächst merkwürdig anmutenden Frage, ob es lange oder kurze Vokale gibt, haben sich die Verfasser aus einem völlig neuen Blickwinkel heraus befasst: Herkömmliche lerntheoretisch begründete Rechtschreibprogramme wären kaum vorstellbar, wenn sie nicht für Dehnungen oder Schärfungen den Zugangshinweis benutzten, das Kind möge darauf achten, ob ein Vokal lang oder kurz ist. So lauten beispielsweise im Marburger Rechtschreibprogramm – und nicht nur dort – die beiden Regeln, die dem Kind zuerst beigebracht werden:

- *Regel 1:* Auf einen lang gesprochenen Selbstlaut folgt nur *ein* Mitlaut.
- *Regel 2:* Auf einen kurz gesprochenen Selbstlaut folgen mindestens *zwei* Mitlaute.

Dem lernenden Kind wird somit abverlangt zu entscheiden, was ein kurzer oder ein langer Vokal ist, denn sonst könnte es diese beiden Regeln ja überhaupt nicht anwenden. Wenn dies dem Kind nicht gelingen will, wird diese Unterscheidungsfähigkeit – ggf. unter Hinzuziehung einer Sprachheiltherapeutin oder Logopädin – nach besten Kräften geübt. Dies gilt als wichtige Komponente der phonologischen Bewusstheit. Viele Kinder besitzen diese Bewusstheit schon vor der Einschulung, andere lernen sie relativ mühelos. Aber es bleibt ein gehöriger Rest, der es einfach nicht zu begreifen scheint – mit entsprechenden Auswirkungen auf die Rechtschreibung, sofern das Kind nicht beizeiten die elegantere Methode des Abspeicherns der richtigen Schreibweise in seinem inneren visuellen Lexikon lernte. Denn: Gute

Rechtschreiber brauchen keine Regeln – sie haben innere Bilder. Zur Beschreibung der inneren Repräsentationen beim Schreiben siehe *Thomé* (1999).

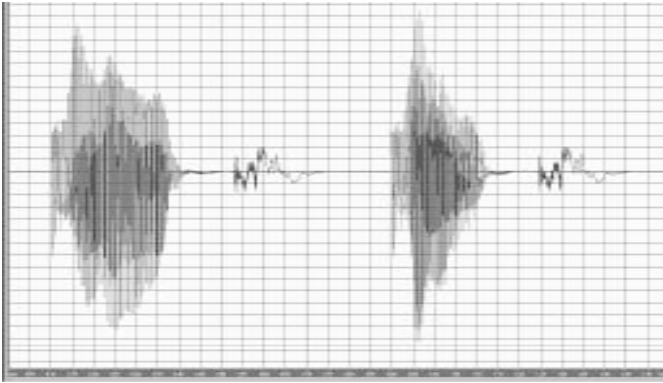
Zurück zu den langen und kurzen Vokalen: Kann von einem Kind tatsächlich erwartet werden, anhand der gehörten Länge eines Wortes zu entscheiden, ob ein Vokal darin lang oder kurz ist? Eigentlich nicht, denn dann müssten wir Menschen einen inneren Schwellenwert besitzen, dessen Zeitdauer sogar objektivierbar sein müsste. Unterhalb dieser Schwelle würde dann ein Vokal als kurz, oberhalb als lang gelten.

Berücksichtigt werden sollte ebenfalls die Tatsache, dass in langsam gesprochenen Sätzen die „kurzen“ Vokale häufig messbar länger sind als die „langen“ Vokale in schnell gesprochenen Sätzen. Schon aus diesem Grunde scheint es entweder kein inneres zeitlich orientiertes Eichnormal für „kurze“ oder „lange“ Vokale zu geben, oder wir benötigen hierfür zusätzliche bzw. andere Informationen.

Diese könnten zum Tragen kommen, wenn argumentiert würde, dass es beim Unterscheiden zwischen kurzen und langen Vokalen nicht um einen Absolutwert, sondern lediglich um einen Relativwert bezogen auf das jeweilige Wort ginge.

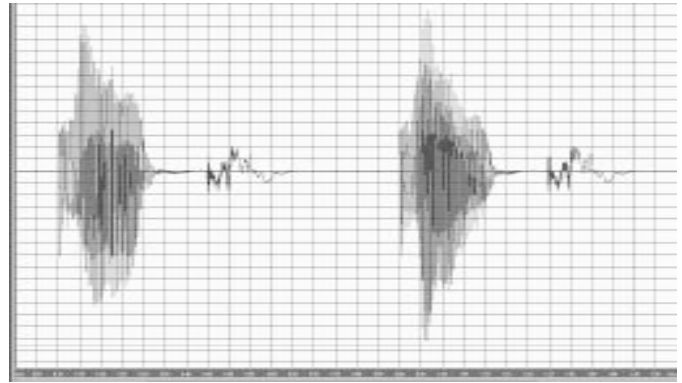
Dem auf diesem Gebiet arbeitenden Wissenschaftler ist natürlich bekannt, dass die absolute messbare Dauer eines Wortsegments tatsächlich unwichtig ist (zumindest für unser gesprochenes Hochdeutsch) und dass es allenfalls der relative Unterschied zwischen zwei Segmentlängen ist, der einen Bedeutungsunterschied begründet und als (Vokal-)Quantität bezeichnet wird.

Bild 1: „Gook“ und „Gock“ im O-Ton



Links „Gook“ mit 630 ms, davon „o“ = 280 ms; rechts „Gock“ mit 550 ms, davon „o“ = 200 ms

Bild 2: „Gook“ elektronisch verkürzt



Links „Gook“ mit dem nunmehr auf 200 ms verkürzten „o“; rechts unverändert das „o“ = 200 ms: Trotz gleicher Länge bleibt der Unterschied hörbar!

Bild 3: (Verkürztes) „o“ aus „Gook“

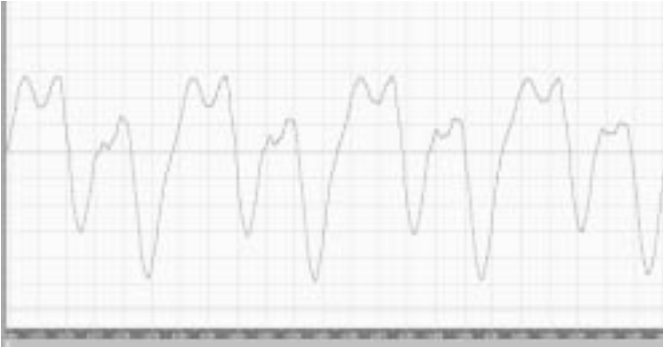
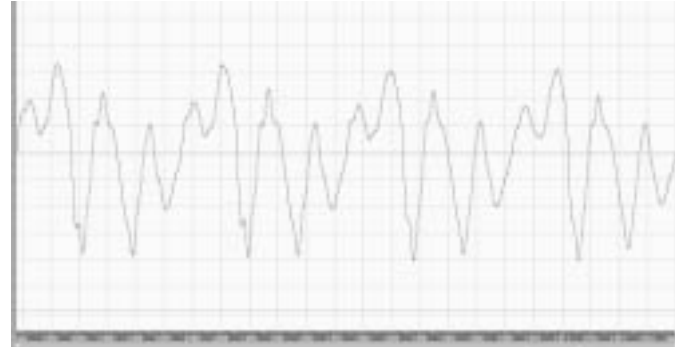


Bild 4: „o“ aus „Gock“



Dass Frequenzunterschiede in diesem Zusammenhang einen ungleich höheren Stellenwert einnehmen, ist ebenfalls in vielen Werken zur Phonetik beschrieben.

Was aber bedeutet das für den Schüler und die Schülerin, die die „langen“ und „kurzen“ Vokale offenbar nicht wahrnehmen können, in Wirklichkeit?

Stellen wir uns einmal ein Wort vor, das wir uns innerlich zweimal vorsprechen – einmal mit kurzem und einmal mit langem Vokal. Genau dieses haben die Verfasser empirisch nachvollzogen. Für die Versuche wurde ein sinnfreies Wort in zwei Varianten gesprochen, die sich tatsächlich nur durch die Länge des Vokals unterschieden, nämlich die Worte „Gook“ und „Gock“. Sinnfreie Worte wurden deswegen gewählt, um einen Rückgriff auf ein vermutlich vorhandenes inneres Lexikon zu vermeiden, mit Hilfe dessen die lang/kurz-Unterscheidung aus der lexikalischen Kenntnis heraus käme und nicht anhand der Hörwahrnehmung.

Das im Folgenden beschriebene Experiment kann von jedem, der über einen PC mit Soundkarte sowie einen Sound-Editor verfügt, ohne große Probleme nachvollzogen werden. Entsprechend ausgestatteten

Schulen sei dies Experiment (vielleicht auch mit anderen Worten aus einer Fremdsprache) für interdisziplinäre Projektarbeit ans Herz gelegt.

Die o.g. beiden sinnfreien, nacheinander mit derselben Grundfrequenz  $f_0 = 170$  Hz Worte „Gook“ und „Gock“ gesprochenen Worte wurden aufgezeichnet.

Die Aufzeichnungen unterschieden sich voneinander auf ersten Anschein tatsächlich nur durch die Länge ihres Vokals: Im Wort „Gook“ ist er 280 ms lang – im Wort „Gock“ sind es nur 200 ms (Bild 1).

Somit ergibt sich eine Gesamtlänge bei „Gook“ von 630 ms, bei „Gock“ von 550 ms. Gutschreibende Probanden jeglichen Alters hatten keine Mühe, die beiden Worte fehlerlos zu identifizieren, ohne die geschriebenen Versionen zu sehen. Hieraus ließe sich zunächst hypothetisch folgern, dass der Längenunterschied von 80 ms hörbar und in die Schriftsprache umsetzbar ist.

Zur Überprüfung dieser Hypothese wurden nun vermittels eines handelsüblichen Sound-Editors<sup>1</sup> 80 ms aus dem „o“ des Wortes „Gook“ herausgeschnitten – natürlich unter Beachtung des Grundsatzes, dass nur im Nulldurchgang eines Schwin-

gungszuges editiert werden darf, um Knackgeräusche zu vermeiden. Das Ergebnis waren nun zwei Worte sowohl mit identischer Vokaldauer von jeweils 200 ms als auch mit identischer Gesamtlänge von 550 ms (Bild 2). Diese Worte wurden den Versuchspersonen abermals vorgespielt mit der Aufforderung, den „langen“ und den „kurzen“ Vokal zu identifizieren.

Ergebnis: Die Erkennungsrate mit entsprechender Umsetzung auf die Schreibweisen „Gook“ und „Gock“ war unverändert gut. Damit steht fest, dass die Unterscheidung nicht auf der Zeitebene stattfinden kann, denn die gehörten Vokale waren jetzt gleich lang.

Worin unterscheiden sich nun die beiden Worte? Nur durch ihren Klang - wissenschaftlich ausgedrückt: die darin enthaltenen Frequenzen. Genau diese zeigt das Oszillogramm von vier typischen Schwingungszügen des „o“ von „Gook“ (Bild 3) sowie des „o“ von „Gock“ (Bild 4).

Selbst ein in dieser Materie ungeübter Leser wird beim Vergleich der beiden Abbildungen den deutlichen Unterschied in den

<sup>1</sup> Verwendet wurde: Cool-Edit 2000, Syntrillium Software Corporation

Wellenformen der beiden „o“ erkennen. Unterschiede in der Schwingungsform bedeuten zugleich aber auch eine unterschiedliche Frequenzstruktur. Die Frequenzanalyse der beiden „o“ brachte die endgültige Bestätigung:

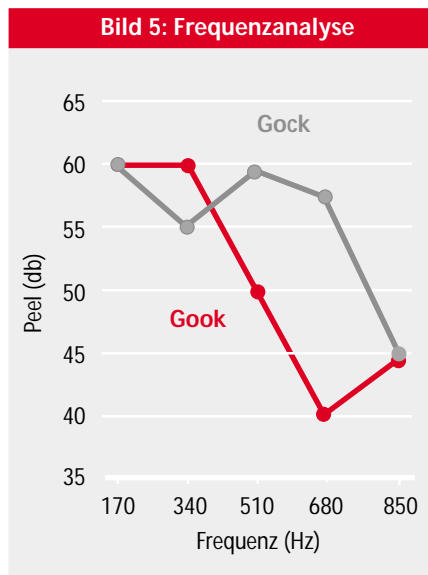
In Bild 5 sind die Grundfrequenz  $f_0$  von 170 Hz sowie die nächsten vier Harmonischen (340, 510, 680 und 850 Hz) mit ihren relativen Pegeln eingetragen. Es ist deutlich zu sehen, dass die Maxima des „o“ in „Gook“ gegenüber dem „o“ in „Gock“ deutlich verschoben sind. Das ist gleichbedeutend mit veränderten Formanten F1 und F2.

Zur Absicherung dieses Experimentes wurden – unter Benutzung desselben Sound-Editors – die abgebildeten Gemische aus fünf Sinustönen in ihren unterschiedlichen Pegeln aufgezeichnet und abgespielt. Dies klingt naturgemäß etwas synthetisch, aber dennoch waren das „lange o“ aus „Gook“ sowie das „kurze o“ aus „Gock“ deutlich zu unterscheiden. Hiermit ist mit relativ einfachen Mitteln bestätigt, was die Phnologie längst weiß, die lerntheoretischen Grundlagen vielerorts aber noch vernachlässigen:

Wir treffen unsere Entscheidung „kurzer oder langer Vokal?“ aufgrund der Wahrnehmung von Frequenzunterschieden, also leicht verschobener Formanten und nicht aufgrund der zeitlichen Länge. Die Erkennung „kurz oder lang?“ ist also nicht ein Problem der zeitlichen Auflösung sondern ein Problem der inneren Frequenzanalyse durch das hörende Kind. Damit hätten Pädagogen, Sprachheillehrer und Logopäden zugleich eine Erklärung dafür, dass der Hinweis „Aber den Längenunterschied musst du doch hören!“ gegenüber jedem Kind solange erfolglos bleiben muss, wie dessen Tonhöhenanalyse im Bereich des zentralen Hörens nicht altersgerecht entwickelt ist.

Der oben geschilderte Versuch kann mit den Vokalen „a“ - „e“ - „i“ - „u“ wiederholt werden. Dies wurde ebenfalls ausprobiert: Mit Ausnahme des Vokals „a“ führten Versuche der Autoren zu vergleichbaren Ergebnissen.

Bestätigung findet dies im Ergebnis einer finnischen Studie (Holopainen et al., 1997), in der bei sprachauffälligen Kindern mit objektiven Messungen nachgewiesen wurde, dass deren Tonhöhendiskrimination im Vergleich zu einer Kontrollgruppe deutlich schwächer war. Eine von der Aufmerksamkeit unabhängige Wellenform auditiver



EEG-Potenziale, die als „Mismatch Negativity“ (MMN) bekannt ist, wurde an Kindern mit sprachlicher Entwicklungsverzögerung und an Kindern mit normaler Sprachentwicklung untersucht. Unter MMN versteht man die Darbietung zweier auditiver Reize, die sich in einer beliebigen Eigenschaft unterscheiden. Sie werden in unterschiedlicher Häufigkeitsverteilung, z.B. 85 % zu 15 % dargeboten. Dabei entsteht aufgrund des seltener auftretenden und damit von der Erwartung abweichenden Reizes automatisch ein negatives EEG-Potential, das als Mismatch Negativity bezeichnet wird. Gemessen wurde u. a. die höchste Amplitude der MMN-Reaktion auf den Frequenzunterschied zwischen 500 und 553 Hz. In 85 % Töne von 500 Hz waren also 15 % Töne von 553 eingestreut. Etwa so: 500 - 500 - 500 - 500 - 553 - 500 - 500 - 500 - 500 - 500 - 553 - 500 - 500 ... usw.

Im Gesamtdurchschnitt war die Amplitude dieses MMN bei den sprachauffälligen Kindern gegenüber den Kontrollkindern signifikant geschwächt. Daraus schlossen die finnischen Forscher, dass sprachauffällige Kinder ein Defizit in der automatischen auditiven Verarbeitung von Frequenzunterschieden aufweisen. Da die MMN-Reaktion die zentrale Hörverarbeitung abbildet und modalitätsspezifisch für auditive Stimuli ist, meinen die Verfasser dieser Studie weiter, dass die MMN-Methode sogar als objektiver Maßstab für die Erhebung zentraler Hörstörungen bei Kindern dienen kann. Dies geschieht inzwischen in Deutschland an vielen Universitätskliniken und (leider) nur bei ganz wenigen HNO-Fachärzten.

## Schlussfolgerung

Die Ergebnisse unseres einfachen Experimentes bestätigen deutlich und für jeden „hörbar“, was schon in Einführungswerken zur Phonetik ausführlich beschrieben wird<sup>2</sup>. Betrachtet man dazu die Ergebnisse der finnischen Studie, wird deutlich, dass es wenig Sinn macht, mit LRS-Kindern weiterhin das Spiel „Den (Längen)Unterschied musst du doch hören!“ zu spielen. Statt dessen sollten auch (und vor allem) Anhänger der herkömmlichen LRS-Programme, für die das Unterscheiden zwischen „langen“ und „kurzen“ Vokalen notwendig ist, zumindest die Tonhöhendiskrimination<sup>3</sup> sprachauffälliger Kinder erheben oder erheben lassen. Es kann nicht schaden, zusätzlich weitere zentrale Hörfunktionen dieser Kinder überprüfen zu lassen. Bei Auffälligkeiten ist vor der Aufnahme jeglicher LRS-Therapie ein Training der Not leidenden zentralen Automatisierungsfunktionen vorzusehen.

**Anm. der Verfasser:** Interessenten können die Dateien <Gook-280-Gock-200.wav> und <GOOK-200-Gock-200.wav>, eingebettet in eine kleine PowerPoint-Präsentation, unter [Fred.Warne@t-online.de](mailto:Fred.Warne@t-online.de) bei Nennung des Stichwortes „Vokale“ kostenlos abrufen. Die Verfasser bieten allen Personen, die sich mit dieser Thematik wissenschaftlich befassen möchten, ihre Zusammenarbeit an.

## Literatur

- Thomé, G. (1999). Orthographieerwerb. Frankfurt/M.: Peter Lang
- Holopainen, I.E.; Korpilahti P.; Juottonen, K.; Lang, H.; Sillanpaa, M. (1997). Attenuated auditory event-related potential (mismatch negativity) in children with developmental dysphasia. *Neuropediatrics* 28 (5), 253-256

## Korrespondenzanschrift

Fred Warnke  
Im Tannengrund 28  
30900 Wedemark  
[Fred.Warne@t-online.de](mailto:Fred.Warne@t-online.de)

<sup>2</sup> Vgl. Petursson, M; Neppert, J. (1996). *Elementarbuch der Phonetik*. Hamburg: Buske

<sup>3</sup> Die Tonhöhendiskrimination und sechs weitere zentrale Funktionen sind messbar mit dem Brain-Boy-Universal, MediTECH Electronic GmbH, Langer Acker 7, 30900 Wedemark, Tel. 0 5130.97 77 8-0, Fax -22, [service@meditech.de](mailto:service@meditech.de)