

Zeitliche Gewichtung bei der Lautheitsintegration

WOLFGANG ELLERMEIER & SABINE SCHRÖDL
Institut für Psychologie, Universität Regensburg, 93040 Regensburg
E-mail:wolfgang.ellermeier@psychologie.uni-regensburg.de

Zusammenfassung

Die zeitliche Integration der Lautheit im Pegel fluktuierender Schalle wird mit einer neuen psychophysischen Methode untersucht. Dazu werden in einem 2IFC-Verfahren jeweils zwei Exemplare weißen Rauschens von je 1 s Dauer dargeboten. Die Pegel der beiden Reize werden in zehn 100-ms-Segmenten unabhängig voneinander nach einer Zufallsverteilung variiert. Aufgabe der Versuchsperson ist es, zu entscheiden, welches der beiden resultierenden Pegelprofile lauter klingt. Aufgrund einiger tausend solcher Urteile werden bedingte psychometrische (COSS-) Funktionen geschätzt, aus denen sich bestimmen läßt in welchem Maße jedes einzelne Zeitsegment zur Gesamtlautheit beiträgt. Die Ergebnisse von sechs Versuchspersonen zeigen große interindividuelle Unterschiede in den Gewichtungsfunktionen. Insgesamt ergibt sich - im Widerspruch zu gängigen Mittelungsmodellen - ein statistisch signifikanter Trend, Reizanfang und -ende stärker zu gewichten als den mittleren Bereich.

Einführung

Die Frage, wie bei Schallen, die im Pegel fluktuieren, ein Lautheitseindruck gebildet wird, ist bisher trotz ihrer offensichtlichen Anwendungsrelevanz wenig erforscht. Die pragmatische Lösung des Problems besteht darin, anzunehmen, der globale Lautheitseindruck eines im Pegel fluktuierenden Schalls entspreche entweder (a) dem eines Schalls konstanten Pegels mit gleicher Energie (Konzept des "Mittelungspegels" (L_{eq} , s. [5]) oder (b) der Lautheit des Pegels, der in einem bestimmten Prozentsatz der Fälle überschritten wird (z.B. N_5 , s. [6], Chap. 16). Beiden Strategien ist gemeinsam, daß davon ausgegangen wird, alle Zeitsegmente würden gleich gewichtet. Deshalb sollen in dieser Untersuchung - in Absetzung von den dargestellten Strategien - nicht *Pegelgewichte* ermittelt werden, die angeben, in welchem Maße hohe, mittlere, und niedrige Pegel das Urteil beeinflussen, sondern *zeitliche Gewichte*, die angeben, wie das auditive System die eingehende Information in Abhängigkeit von zeitlichen Parametern integriert.

Herkömmliche Methoden - etwa der direkten Einschätzung "instantaner" Lautheitseindrücke - sind (1) wegen ihrer Transparenz, (2) wegen möglicher Urteilsartefakte und (3) v.a. weil sie nicht die zeitliche Auflösung bieten, bis in den Sekundenbereich vorzudringen, für die Untersuchung dieser Fragestellung ungeeignet. Eine Alternative bietet sich in dem von Berg [1] vorgeschlagenen Verfahren der "COSS-Analyse". Dabei werden die Entscheidungen eines Beobachters bei einer Wahrnehmungsaufgabe, die mehrere Informationskanäle anspricht, bezüglich des Beitrags der einzelnen Kanäle analysiert, *ohne Einzelurteile über die Konstituenten der Entscheidung abzufragen*. Der experimentelle Kunstgriff, der dies ermöglicht, besteht darin, jede Reizkomponente mit einer unabhängigen, zufälligen Fluktuation (x_i) zu versehen und den auf diese Weise "verrauschten" Gesamtreiz zur Beurteilung darzubieten. Wird nun die Entscheidung des Beobachters als Funktion der Zufallsfluk-

tuationen x_i separat für jeden der beteiligten Informationskanäle analysiert, so ergibt sich jeweils eine psychometrische Funktion (*COSS function*, s.u.), aus deren Varianz sich der Beitrag des betrachteten Informationskanals zur Gesamtentscheidung ablesen läßt. Aufgrund solcher COSS-Funktionen lassen sich Gewichte a_i schätzen, die angeben, wie stark und in welcher Richtung die jeweilige Reizkomponente die Entscheidung beeinflusst.

Bisher wurden v.a. verschiedene Frequenzkanäle als die Reizkomponenten aufgefaßt, deren Beitrag es zu analysieren gilt (z.B. [2,3]). In dieser Untersuchung werden nun verschiedene *zeitliche* Segmente als Informationskanäle definiert und in ihrem Beitrag zur Gesamtlautheit analysiert.

Methode

Mit Hilfe einer Signalprozessorkarte (Tucker-Davis-Technologies [TDT] AP2) wird weißes Rauschen von 1 s Dauer generiert und über einen 16-bit Digital-Analog-Wandler (TDT DD1) mit 50 kHz Samplingrate ausgegeben. Anschließend wird das Signal alle 100 ms stufenförmig durch einen programmierbaren Abschwächer (TDT PA4) in seiner Intensität verändert (s. Abb. 1). In jedem Versuchsdurchgang wurden zwei solche Pegelverläufe dargeboten. Einer ist (i.S. der Signalentdeckungstheorie) als "Rauschen" definiert, der andere als "Signal". Für das Rauschen werden - *unabhängig* in jedem der 10 Zeitsegmente - Schallpegel aus einer Normalverteilung mit Erwartungswert (horizontale Linien in Abb. 1) $\mu_R = 60$ dB SPL und Standardabweichung $\sigma = 2$ dB gezogen. Für das "Signal" beträgt der Erwartungswert $\mu_S = 61$ dB SPL, bei gleicher Standardabweichung. Die Schalle werden diotisch über Kopfhörer (Sony MDR-CD370) dargeboten.

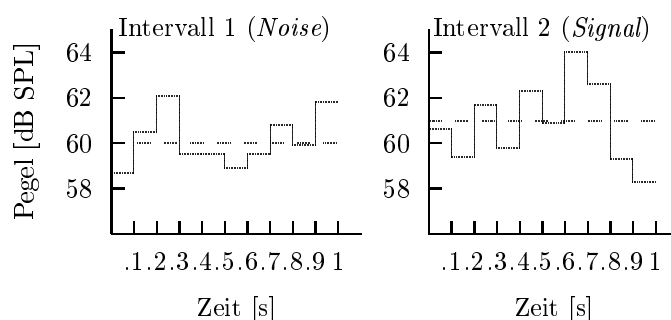


Abbildung 1: Schematische Darstellung eines Versuchsdurchgangs in der verwendeten Forced-Choice-Aufgabe. Es wird weißes Rauschen dargeboten, dessen Pegel zufällig variiert. In diesem Fall enthält das zweite Beobachtungsintervall das "Signal", d.h. Ziehungen aus der Pegelverteilung mit dem höheren Erwartungswert.

Aufgabe der Versuchsperson ist es, ein globales Urteil darüber abzugeben, welches der beiden Pegelprofile eines Versuchsdurchgangs lauter klingt. Insgesamt werden pro Versuchsperson in ca. 10 Sitzungen 3000 solcher forced-choice trials erhoben. Sechs normalhörende Versuchsperso-

nen (einschließlich der Autoren) nahmen an den Experimenten teil.

Ergebnisse

Um zu bestimmen, in welchem Maße ein bestimmtes Zeitsegment das Gesamturteil beeinflusst, wurde die Differenz der Zufallsperturbationen im 1. und 2. Beobachtungsintervall gebildet: $x_{i1} - x_{i2}$. Trägt man nun den Anteil der "Intervall 1"-Antworten als Funktion dieser Differenz ab, so erhält man eine sog. *COSS-Funktion* (COnditional on a Single Stimulus), d.h. das Urteil als Funktion eines einzigen Reizattributs (etwa des Pegels im 2. Zeitsegment, $i = 2$). Insgesamt wurden pro Versuchsperson 20 (2 Beobachtungsintervalle \times 10 Zeitsegmente) solcher COSS-Funktionen geschätzt. Diese wurden in fast allen Fällen hinreichend gut durch kumulative Normalverteilungen beschrieben. Abb. 2 zeigt beispielhaft die Daten einer einzelnen Versuchsperson für eines der zehn Zeitsegmente, zusammen mit den bestangepaßten "normalen Ogiven".

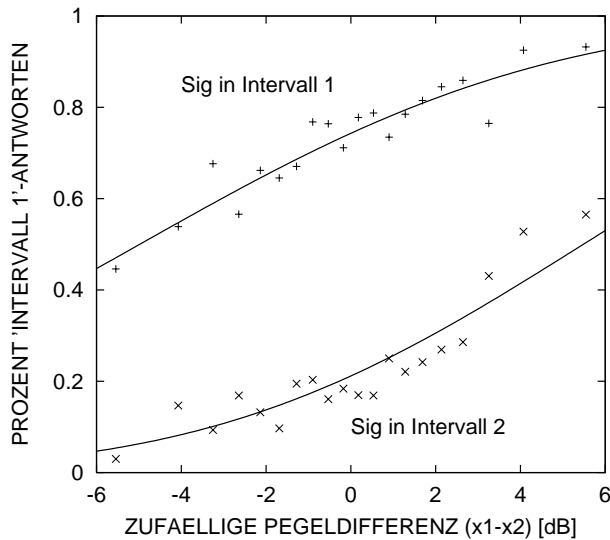


Abbildung 2: Beispiel für eine COSS-Funktion (Versuchsperson PD im zweiten 100-ms-Abschnitt). Dargestellt ist der Anteil von "Intervall 1"-Antworten als Funktion der Differenz der zufälligen Pegelfluktuationen.

Wie in [1] gezeigt, lassen sich aus der Varianz der COSS-Funktionen die Gewichte a_i schätzen, in unserem Falle also quantitative Maße dafür, welchen Einfluß das jeweilige Zeitsegment auf das Gesamturteil hat. "Steile" COSS-Funktionen (wie in Abb. 2) implizieren eine starke Gewichtung, "flache" COSS-Funktionen (genauer: solche mit großer Varianz) bedeuten, daß das jeweilige Zeitsegment nicht berücksichtigt wird (die Differenz $x_{i1} - x_{i2}$ hat keinen Einfluß auf das Urteil).

Abbildung 3 zeigt die mittleren, normierten Gewichte aller 6 Versuchspersonen als Funktion der zehn Zeitsegmente. Der Verlauf weicht systematisch ab von der Annahme "idealer" Gewichte $a_i = 1/n = 1/10$ (horizontale Linie in Abb. 3), wonach alle Zeitsegmente in gleicher Weise berücksichtigt würden. Vielmehr scheint eine Akzentuierung des Reizanfangs vorzuliegen. Mit zunehmender Reizdauer nehmen die Gewichte für fast alle Versuchspersonen systematisch ab. Das zeigt sich auch in einer zweifaktoriellen (Zeitsegmente \times Beobachtungsintervalle) Varianzana-

lyse, bei der lediglich der Zeitfaktor statistisch signifikant wird, $F(9, 45) = 4.43$; $p < 0.001$.

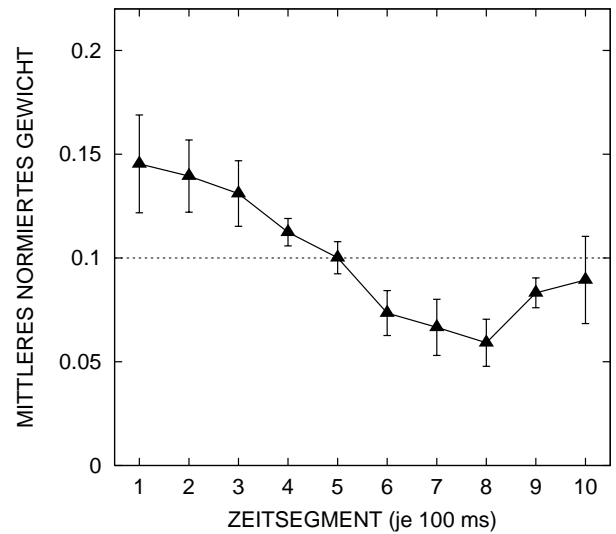


Abbildung 3: Mittlere Gewichte von 6 Vpn. Dargestellt sind die Mittelwerte der auf $\sum a_i = 1.0$ normierten Gewichte als Funktion des jeweiligen Zeitsegments. Die Standardfehler kennzeichnen die interindividuelle Varianz der Gewichte; die horizontale Linie markiert "ideale" Gewichte von $1/10$.

Schlußfolgerungen

Die Ergebnisse zeigen, daß die instantanen Pegel kurz dauernder, nichtstationärer Schalle bei der Bildung der globalen Lautheit nicht gleich gewichtet werden. Diese Ergebnis steht im Widerspruch zu gängigen Algorithmen der Lautheitsberechnung, die in der Regel keine zeitlichen Gewichte vorsehen, entspricht aber Hinweisen aus neueren psychoakustischen Untersuchungen [4]. Weitere Experimente in unserem Labor sollen die Effekte der Reizdauer und der Anzahl der Zeitsegmente aufklären, um zu eruieren, welche sensorischen und kognitiven Mechanismen bei der Bildung der Gesamtlautheit beteiligt sind.

Literatur

- [1] Berg, B.G. (1989). Analysis of weights in multiple observation tasks. *Journal of the Acoustical Society of America*, 86, 1743-1746.
- [2] Berg, B.G. & Green, D.M. (1990). Spectral weights in profile listening. *Journal of the Acoustical Society of America*, 88, 758-766.
- [3] Ellermeier, W. & Berg, B.G. (1996). Perceptual contrast enhancement in profile analysis tasks. In A. Schick & M. Klatte (Eds.), *Contributions to the 7th Oldenburg Symposium on Psychoacoustics* (pp. 405-416). Oldenburg: BIS.
- [4] Patterson, R.D. & Irino, T. (1998). Modeling temporal asymmetry in the auditory system. *Journal of the Acoustical Society of America*, 104, 2967-2979.
- [5] Schick, A. (1990). *Schallbewertung: Grundlagen der Lärmforschung*. Berlin: Springer.
- [6] Zwicker, E. & Fastl, H. (1999). *Psychoacoustics: Facts and models* (2. Auflage). Berlin: Springer.