Gehörbezogene Modellierung der Dissonanz

Anna Rieger[1], Arne Oetjen[2], Steven van de Par[2], Hans-Peter Rabl[1]
[1] *OTH Regensburg, Labor für Entwicklungsanwendungen, 93053 Regensburg, Deutschland,*
Email: anna.rieger@st.othr.de, anna.rieger@uol.de
[2] *CvO Universität Oldenburg, Abt. Akustik, Cluster of Excellence "Hearing4All", 26129 Oldenburg, Deutschland,*
Email: arne.oetjen@uol.de

Einleitung

Die Begriffe Konsonanz und Dissonanz sind vor allem bekannt aus der (westlichen) Musiktheorie. Dabei ist die Dissonanz als Phänomen definiert, bei welchem gleichzeitig dargebotene Töne einen eher unangenehmen Höreindruck hervorrufen können [1]. Die Konsonanz wird für gewöhnlich als das direkte Gegenteil der Dissonanz beschrieben [2].

Eine bekannte Theorie für das Dissonanz-Phänomen ist das Vorhandensein von Schwebefrequenzkomponenten, die durch nahe aneinanderliegende Teiltöne erzeugt werden [3–6]. Nicht zu verwechseln ist dieses Phänomen mit dem psychoakustischen Parameter Rauigkeit, der nach Untersuchungen in [7] die Dissonanzwahrnehmung nicht beschreiben kann.

Während in der Musik und in bekannten Stimmungssystemen der Konsonanz- und Dissonanzgrad für bestimmte Tonintervalle definiert ist, ist das Konzept für Umweltgeräusche jedoch nur limitiert anwendbar, da auftretende beliebige Frequenzintervalle und natürlich vorkommende stark variierende Pegelverhältnisse davon nicht berücksichtigt werden.

Ein aktuelles Beispiel sind elektrisch angetriebene Fahrzeuge, die durch Komponenten wie Elektromotoren, Inverter, Getriebe oder auch die gesetzlich vorgeschriebenen AVAS (Acoustic Vehicle Alerting System) prominente Tonkombinationen erzeugen, welche eine Auswirkung auf die generelle Geräuschqualität haben [8–10].

Ziel des im Folgenden beschriebenen Hörexperiments war es, die menschliche Empfindung von Dissonanz und genereller Präferenz nicht-musikalischer Geräusche quantitativ zu bestimmen und daraus ein gehörgerechtes Berechnungsmodell abzuleiten, das unter anderem im Bereich des Sound-Design eingesetzt werden kann.

Experiment: Dissonanz vs. Präferenz

Hinsichtlich ausgeprägter Dissonanzen liegt der Ansatz nahe, Frequenzverhältnisse abzuändern und so wiederrum Konsonanzen zu erzeugen. Bei technischen Systemen und entsprechend fortgeschrittenen Entwicklungsständen ist dies nur eingeschränkt möglich, da charakteristische Frequenzen ohne aufwendige konstruktive Adaptionen nicht oder nur schwer verändert werden können. Eine gängige Methode ist daher die Adaption spezifischer Pegel einzelner Komponenten, um die Geräuschqualität zu verbessern.

Bislang war der Einfluss von Pegelverhältnissen zwischen Teiltönen auf die empfundene Dissonanz kein Gegenstand laufender Untersuchungen.

Bei einer Veränderung eines Pegelverhältnisses zweier Töne zeigt sich eine Änderung im Einhüllendenbereich der jeweiligen Tonkombination. Das Experiment unterliegt somit der Hypothese, dass die Dissonanz durch eine Änderung der Pegelverhältnisse von Teiltönen in Geräuschen beeinflusst werden kann.

Stimuli und Methode

Die synthetisch erzeugten Stimuli haben sich aus jeweils zwei harmonischen Tonreihen und einem bandpassgefiltertem Hintergrundrauschen mit fest definiertem SNR=15 dB zusammengesetzt. Eine harmonische Reihe besaß dabei jeweils drei Teiltöne, wovon die Obertöne einen Abfall im Schalldruckpegel von $-6\,\mathrm{dB}$ aufwiesen. Neben der ersten harmonischen Reihe, die für alle Geräusche identisch war ($f_0 = C6 = 1046,5\,\mathrm{Hz}$, $2 \cdot f_0$, $3 \cdot f_0$), war die zweite harmonische Reihe um ein Frequenzverhältnis $\mathbf{x}$ zur ersten Reihe verschoben ($\mathbf{x} \cdot f_0$, $2 \cdot \mathbf{x} \cdot f_0$, $3 \cdot \mathbf{x} \cdot f_0$). Die experimentellen adaptiven Variablen waren dabei die Pegel, der um das Frequenzverhältnis $\mathbf{x}$ verschobenen Teiltöne. Der höhere Frequenzbereich und eine Interaktion von harmonischen Tonreihen wurde gewählt, da beispielsweise elektrisch angetriebene Fahrzeuge ab eben diesem Bereich prominente komplexe Tonreihen emittieren [9, 11].

Als Versuchsmethode wurde eine adaptive 2-Intervall AFC (Alternative Forced Choice) Methode eingesetzt, bei welcher generell immer ein *Test*geräusch mit einem *Referenz*geräusch verglichen wird [12–14]. Dabei haben Testpersonen entweder „Which interval sounds more dissonant?" (*Welches Geräusch hört sich mehr dissonant an?*) oder „Which interval would you prefer as a vehicle sound?" (*Welches Geräusch würden Sie mehr als Fahrzeuggeräusch präferieren?*) beantwortet.

Bei der Entscheidung der Testpersonen für eines der beiden Geräusche wurden im Programmhintergrund die Teiltonpegel schrittweise reduziert, bis Testpersonen am Punkt subjektiver Gleichheit (PSE) angekommen waren. An diesem Punkt herrscht eine 50 % Wahrscheinlichkeit, welches der beiden Geräusche nun noch als mehr dissonant empfunden oder mehr präferiert wird. Als Ergebnis haben Testpersonen demnach eine bestimmte Pegeländerung von Teiltönen vorgenommen, um die empfundene Dissonanz oder Präferenz zweier unterschiedlicher Geräusche aneinander anzupassen.

Tabelle 1 zeigt die genau eingestellten Frequenzintervalle $\mathbf{x}$ als Dezimalwert (gleichstufige Stimmung), die Einordnung aus der Musik, ob Intervalle als eher konsonant oder dissonant bekannt sind, als auch die Deklaration, ob es sich um ein *Test*- bzw. *Referenz*geräusch gehandelt hat. Als *Referenz*geräusche wurden jene Intervalle aus-

gewählt, die in der Musik als eher konsonant bekannt sind. Als *Test*geräusche hingegen wurden als dissonant eingestufte Intervalle präsentiert, um Vergleiche zwischen konsonant und dissonant eingestuften Geräuschen zu erzielen. Eine Ausnahme bildeten die *Test*geräusche *+/-Quarte*, um ebenso Vergleiche von Geräuschen mit unterschiedlichen Konsonanzgraden im Experiment darzubieten.[1]

Tabelle 1: Übersicht und Zuordnung eines Teils der im Hörversuch verwendeten Frequenzintervalle x.

Frequenzintervall **x**			
Name	Wert	Grad	Art
-Septime	0,561	dissonant	Test
-Tritonus	0,707	dissonant	Test
-Quarte	0,749	konsonant	Test
-Sekunde	0,944	dissonant	Test
+Sekunde	1,059	dissonant	Test
+Quarte	1,335	konsonant	Test
+Tritonus	1,414	dissonant	Test
+Septime	1,782	dissonant	Test
+Terz	1,260	konsonant	Referenz

Ein Teil der Geräusche besaß eine zweite harmonische Tonreihe, welche zu höheren Frequenzen hin verschoben wurde ($x > 1$). Wiederrum waren auch Geräuschbeispiele enthalten, bei denen die zweite Tonreihe zu niedrigeren Frequenzen hin verschoben wurde ($x < 1$).
Im Hörversuch wurden die Testgeräusche (*-Septime* bis *+Septime*) mit der Referenz *+Terz* verglichen. Die Dissonanz als auch die Präferenz wurden in randomisierter Reihenfolge abgefragt, ebenso wurde die Wiedergabe der jeweiligen Vergleiche randomisiert.
Die Wiedergabe der Geräusche erfolgte diotisch über Kopfhörer (Sennheiser HD 650) an einer externen Soundkarte in einwandigen Hörkabinen. Der Hörversuch wurde anhand des Referenzstimulus *+Terz* auf 65 dB kalibriert.

Ergebnisse: Dissonanz vs. Präferenz

Abbildung 1 zeigt die gemittelten Versuchsergebnisse mit Standardfehlern von 23 Versuchspersonen (13m, 10w, im Alter von 20-44 Jahren). Gezeigt sind alle Vergleiche der Testgeräusche mit der Referenz *+Terz*.[2] Positiv aufgetragene Werte signalisieren eine Änderung der Teiltonpegel der Testgeräusche, wohingegen negative Werte bedeuten, dass Teiltonpegel des Referenzgeräusches reduziert wurden.
Aus den Ergebnissen lässt sich schlussfolgern, dass sich die Dissonanz und Präferenzurteile signifikant voneinander unterscheiden, Menschen somit für diese Art von Geräuschen die Dissonanz nicht direkt mit der generellen empfundenen Präferenz gleichsetzen.
Außerdem können bekannte Theorien aus der Musik beobachtet werden. Als dissonant bekannte Intervalle (*Septimen, Tritoni und Sekunden*) benötigten am meisten Pegelreduktion, um gleich dissonant zu einer Referenz empfunden zu werden. Für als konsonant bekannte Intervalle (hier: *Quarten*) musste nur wenig Pegelreduktion vorgenommen werden, um die emfpundene Dissonanz anzugleichen.

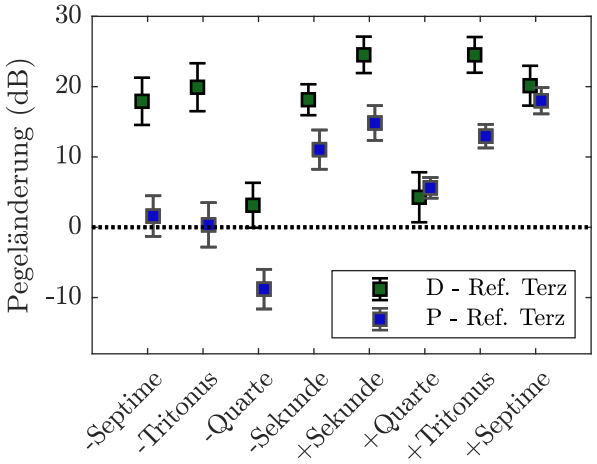

Abbildung 1: Gemittelte Ergebnisse der Dissonanz- und Präferenzurteile inklusive Standardfehler für 23 Versuchspersonen. In grün: Dissonanz; In blau: Präferenz;

Die Dissonanz wurde außerdem für die Intervalle (*-Septime/+Septime*), (*-Tritonus/+Tritonus*) und (*-Sekunde/+Sekunde*) jeweils mit sehr ähnlicher Pegeländerung eingestellt, obwohl die zweite harmonische Reihe einmal zu höheren Frequenzen und einmal zu tieferen Frequenzen hin verschoben wurde. Nach Erkenntnissen aus der Musik wäre das auch zu erwarten gewesen, da das sich ausbildende Frequenzintervall x gleich ist (vgl. z.B. *-Septime*=0,561 und *+Septime*=$\frac{1}{0{,}561} = 1{,}782$).
Vergleicht man für eben diese Datenpunkte die eingestellten Pegeländerungen für die Präferenzurteile fällt auf, dass dissonante Frequenzintervalle mit $x < 1$ (*-Septime*, *-Tritonus*) nahezu keine Pegeländerung notwendig hatten, um gleich präferiert zu werden zum hier konsonant eingestuften Referenzgeräusch *+Terz*. Für beispielsweise *+Septime* und *+Tritonus* waren steigende Pegeländerungen notwendig, um die Präferenz zum Refrenzgeräusch einzustellen.

Modellierung: Dissonanz

Zur gehörgerechten Modellierung der Dissonanz für technische Geräusche wurde ein Modellansatz gewählt, der im ersten Teil eine gehörgerechte Vorverarbeitung anwendet und in zweiten Teil eine Merkmalsextraktion über Einhüllendenschwankungen vorsieht.
Das Modell weist die folgenden, wichtigsten Verarbeitungsschritte auf: Ein beliebiges Schalldrucksignal $p(t)$ wird zunächst mittels einer Gammatonfilterbank zur Simulation der Basilarmembran verarbeitet. Für jeden Filterkanal wird danach eine Halbwellengleichrichtung mit anschließender Bildung der Hilberteinhüllenden durchgeführt und im nächsten Schritt wie in [15] motiviert eine filterkanalabhängige Autokorrelationsfunktion angewendet. Zur Extraktion der Schwebefrequenzkomponenten werden kanalabhängige Bandpassfilter eingesetzt und an-

[1] Ein Teil der im Experiment präsentierten Stimuli und Versuchsergebnisse sind in dieser Veröffentlichung nicht gezeigt.
[2] Nicht dargestellte Ergebnisse lassen gleiche Schlussfolgerungen zu.

schließend eine Summation über alle Gehörbänder durchgeführt. Eine anschließende Transformation und Summation führt schließlich zu einer Einzahlgröße D, welche als Dissonanz-Indikator gewertet wird. Eine zusätzliche zeitliche Fensterung der gefilterten Zeitsignale ermöglicht außerdem die Auswertung transienter Vorgänge und die Ausgabe der Dissonanz über die Zeit $D(t)$.

Ergebnisse: Dissonanz-Modell

Zum Test des Modells wurde das AFC-Versuchsverfahren so programmiert, dass das Modell anstatt einer Versuchsperson das Experiment durchführt und der entwickelte Dissonanz-Indikator D als Schwellwert zur Beantwortung der Geräuschvergleiche dient. Abbildung 2 zeigt den Vergleich der gemittelten Ergebnisse der Versuchspersonen (grün) sowie das Mittel von zehn Modelldurchläufen für Dissonanz (pink) mit dem Referenzgeräusch *+Terz*.

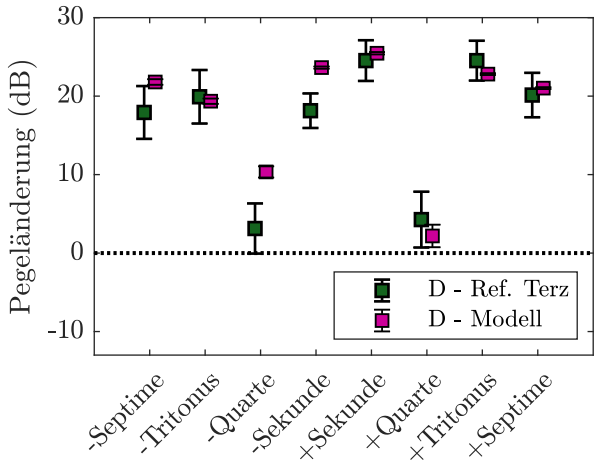

Abbildung 2: Vergleich der Dissonanz-Modellberechnungen (pink) und den gemittelten Ergebnissen des Versuchskollektivs (grün).

Wie der Abbildung zu entnehmen ist, ist das Gehörmodell bereits fähig die empfundene Dissonanz sehr ähnlich im Vergleich zum Mittel des Versuchskollektivs zu bewerten. Hervorzuheben ist dabei, dass das Modell anstelle symbolischer Inputs (Vorgabe einzelner Frequenzen) beliebige Schalldrucksignale (auch transient) $p(t)$ verarbeiten kann. Mithilfe der Anwendung gehörgerechter Vorverarbeitung kann außerdem sichergestellt werden, dass die Geräuschinformation empfindungsgerecht für die Merkmalsextraktion bereitgestellt wird. Das Modell beruht zudem auf keinerlei a priori Wissen zu Konsonanz- oder Dissonanz-Theorien.

Neben den im Hörversuch verwendeteten Frequenzintervallen wurde zusätzlich eine Modellberechnung gestartet, die viele verschiedene, auch nicht abgetestete Frequenzintervalle bewertet hat. Hierfür wurden wieder Geräusche mit zwei harmonischen Reihen erstellt, die Intervalle zwischen *-Oktave:* $x = 0,5$ und *+Oktave:* $x = 2$ enthielten. Abbildung 3 zeigt den Dissonanz-Verlauf der Modellberechnung für beliebige Frequenzintervalle in feinem Raster. Dabei lässt sich ein charakteristischer Dissonanz-Verlauf beobachten, der die aus der Musik bekannten Tonintervalle (farblich markiert) ähnlich der aus der Musiktheorie bekannten Paradigmen bewerten kann.

Zusammenfassung

Im Hinblick auf effektives Sound-Design bei multitonalen Geräuschen am beispiel von elektrisch angetriebenen Fahrzeugen ergibt sich das Problem der Konsonanz- bzw. Dissonanzbeurteilung. Durch künstlich erzeugte Geräusche oder beispielsweise Stromformung können die Geräusche aktiv beeinflusst werden, jedoch bieten etablierte psychoakustische Parameter bislang keine geeignete Optimierungsgröße hierfür.

Die Konsonanz und Dissonanz ist bekannterweise stark abhängig vom jeweiligen Frequenzintervall, da bei

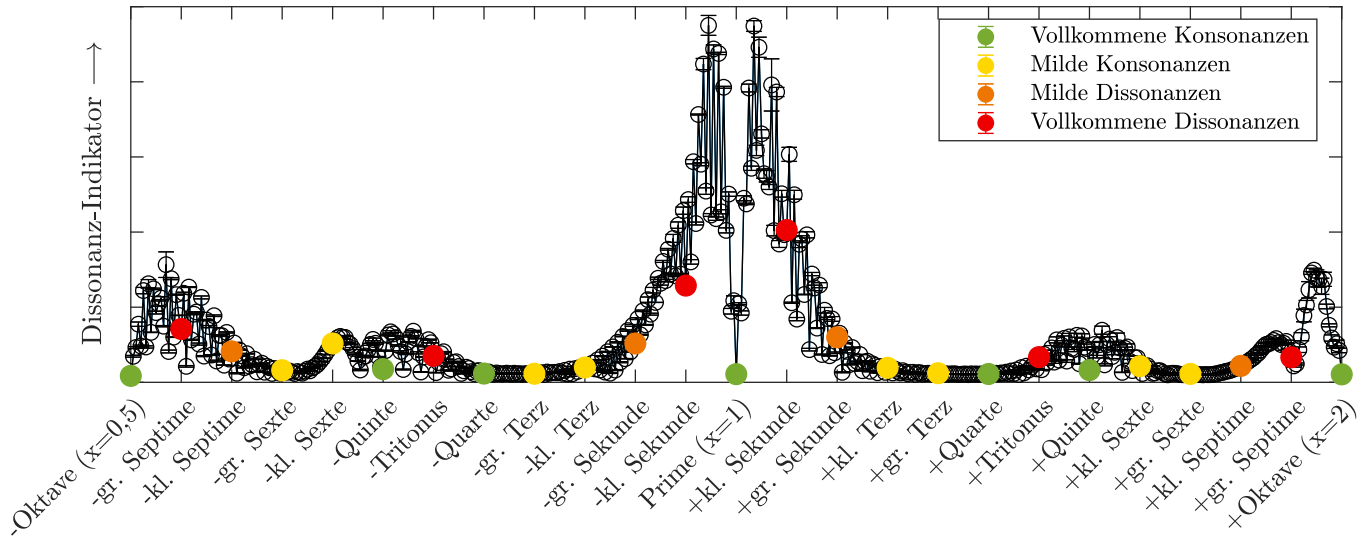

Abbildung 3: Modellberechnung für beliebige Frequenzintervalle von x =*-Oktave* bis x =*+Oktave*. Aus der Musik bekannte Intervalle sind entsprechend des jeweiligen angenommenen Konsonanz- und Dissonanzgrades farblich gekennzeichnet.

technischen Systemen der Frequenzgehalt oft nicht uneingeschränkt verändert werden kann, sollte der Einfluss von Tonpegelverhältnissen auf die Dissonanzwahrnehmung untersucht werden.
Es hat sich herausgestellt, dass die Dissonanz von Geräuschen durch die Justierung von Tonpegeln beeinflusst werden kann. Als sehr dissonant emfpundene Frequenzintervalle können somit alleinig durch eine Veränderung von Teiltonpegeln als konsonanter wahrgenommen werden. Dissonanz und Präferenz werden von Zuhörenden unterschiedlich wahrgenommen. In weiterführenden Analysen hat sich außerdem gezeigt, dass die Dissonanz der für die Wahrnehmung entscheidende Parameter war.
Tendenzen hinsichtlich der musikalischen Einordnung konsonanter und dissonanter Ton-Intervalle konnten durch den Hörversuch ebenso bestätigt werden.
Zur Modellierung der Dissonanz wurde ein gehörbezogener Ansatz herangezogen. Zusammengefasst besteht der Modell-Ansatz aus einer gehörgerechten Signalvorberarbeitung sowie einer Merkmals-Extraktion für die letztliche Dissonanzberechnung. Aus dem Einhüllendenbereich der Geräusche werden filterkanalabhängig Informationen extrahiert, welche summiert über alle Kanäle als Schlüsselmerkmale benutzt werden. Durch eine erste Optimierung und einem Vergleich von Modell- zu Versuchsdaten stellte sich heraus, dass der Modellansatz die Subjektivurteile zufriedenstellend nachbilden kann.
Somit konnte ein erster vielversprechender Modellansatz zur Berechnung der empfundenen Dissonanz entwickelt werden, der:

- beliebige Schalldrucksignale p(t) verarbeiten
- ebenso transiente Vorgänge berücksichtigen
- gehörgerecht und ohne a priori Wissen multi-tonale Geräusche bezüglich der empfundenen Dissonanz bewerten

kann.

Ausblick

In weiteren Untersuchungen soll die Datenbasis zu Subjektivurteilen empfundener Dissonanz und allgemeiner Präferenz ausgeweitet werden. Darunter sind Experimente geplant, die verschiedenste Signalparametervariationen abdecken und dabei synthetisch generierte als auch reale Prüfstands- oder Fahrzeuggeräusche miteinbeziehen.
Indes ist eine kontinuierliche Validierung sowie Weiterentwicklung des grundlegenden Dissonanz-Modells geplant.
Zusätzlich dazu soll der Einfluss der Dissonanz auf die allgemeine Geräuschqualität modelliert werden.

Literatur

[1] ASA Standards: dissonance. URL: `https://asastandards.org/Terms/dissonance/`.

[2] ASA Standards: consonance. URL: `https://asastandards.org/Terms/consonance/`.

[3] Von Helmholtz, H.: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Springer Fachmedien Wiesbaden, 6. Auflage, 1913.

[4] Plomp, R. und Levelt, W. J. M.: Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America*, 38:548–560, 1965.

[5] Sethares, W.: Local consonance and the relationship between timbre and scale. *The Journal of the Acoustical Society of America*, 94:1218–1228, 1993.

[6] Terhardt, E. (Hrsg.): *Akustische Kommunikation: Grundlagen mit Hörbeispielen*. Springer-Verlag Berlin Heidelberg, 1998.

[7] Reffgen, M. und Fiebig, A.: Die Grenzen der Rauigkeit zur Beschreibung der Dissonanzwahrnehmung. *Fortschritte der Akustik - AIA-DAGA 2013*, 391–394.

[8] Bell, L. H. und Bell, D. H. (Hrsg.): *Industrial Noise Control: Fundamentals and Applications*. Marcel Dekker, Inc., New York, 2. Auflage, 1993.

[9] Fingerhuth, S.: *Tonalness and Consonance of Technical Sounds*. Dissertation, RWTH Aachen, 2009.

[10] Parizet, E. und Fingerhuth, S.: Subjective evaluation of tonalness and relation between tonalness and unpleasantness. *Proceedings Internoise 2009*.

[11] Tschöke, H. et al. (Hrsg.): *Elektrifizierung des Antriebsstrangs: Grundlagen - vom Mikro-Hybrid zum vollelektrischen Antrieb*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2019.

[12] Kingdom, F. und Prins, N. (Hrsg.): *Psychophysics: A practical introduction*. Elsevier/Academic Press, Amsterdam, 2. Auflage, 2016.

[13] Ewert, S.: AFC - A modular framework for running psychoacoustic experiments and computational perception models. *Fortschritte der Akustik - AIA-DAGA 2013*, 1326–1329.

[14] Ewert, S.: AFC: The psychophysical-measurement package for MATLAB. URL: `http://medi.uni-oldenburg.de/afc/index.htm`.

[15] Meddis, R. und O'Mard, L.: A unitary model of pitch perception. *The Journal of the Acoustical Society of America*, 102(3):1811–1820, 1997.