

Entmaskierung durch binaurale und harmonische Signalmerkmale von Tonkomplexen

M. Klein-Hennig, M. Dietz, V. Hohmann

Medizinische Physik, Carl-von-Ossietzky Universität, 26121 Oldenburg, Germany, Email: martin.klein.hennig@uni-oldenburg.de

Einleitung

Auditorische Szenenanalyse erlaubt es Menschen, in komplexen akustischen Szenen mehrere Geräuschquellen zu identifizieren, lokalisieren und ihrem zeitlichen Verlauf zu folgen. Dazu werden mehrere verfügbare Signaleigenschaften extrahiert und kombiniert (siehe z.B. [1]). Binaurale Merkmale wie die interaurale Laufzeitdifferenz (ITD, IPD) oder Pegeldifferenz (ILD) werden zur Lokalisation verwendet. Die Tonhöhe eines Signals erleichtert die Trennung von Signalen aus anderen Quellen. Für die auditorisch motivierte Modellierung der Szenenanalyse ist Wissen über die Kombination und Verarbeitungsreihenfolge dieser Signalmerkmale erforderlich.

Experimente, die sich mit der Wahrnehmung von Harmonizität beschäftigen, verwenden häufig harmonische Tonkomplexe, deren Komponenten ganzzahlige Vielfache der Grundfrequenz F_0 des Komplexes sind. Wird eine Komponente in einem solchen Tonkomplex verstimmt, so dass sie kein ganzzahliges Vielfaches der Grundfrequenz mehr ist, wird sie salienter und scheint aus dem Tonkomplex herauszustechen (siehe z.B. [2]). In dieser Studie wurde daher Verstimmung eingesetzt, um den Einfluss harmonischer Signalmerkmale auf die Detektion einer Komponente eines Tonkomplexes zu untersuchen. Im Kontext eines Detektionsexperiments kann sich der Gewinn durch Verstimmung als "Mistuning Masking Level Difference" (MMLD) bezeichnen lassen.

Durch binaurale Signalmerkmale lassen sich Mithörschwellen ebenfalls senken. Dies kann z.B. durch das Aufprägen eines interauralen Phasunterschieds auf das Testsignal realisiert werden. Die Schwellensenkung durch binaurale Merkmale wird als "binaural masking level difference" (BMLD, siehe z.B. [3]) bezeichnet. In dieser Studie wird BMLD als binauraler Effekt verwendet und seine Kombination mit MMLD untersucht.

Stimuli und Methoden

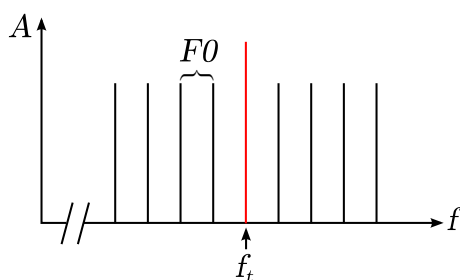


Abbildung 1: Schema des Stimulus-Spektrums

Die verwendeten Stimuli waren harmonische Tonkom-

plexe, generiert durch die Aufsummierung mehrerer Reintöne in zufälliger Phase. Abbildung 1 zeigt ein schematisches Spektrum der Stimuli. Sie bestanden aus einem Maskierer und einem Testsignal: der Maskierer war ein harmonischer Tonkomplex der Grundfrequenz F_0 mit 8 Komponenten, die um die Frequenz f_t des Testsignals zentriert waren (s. Abb.1). Jede Komponente hatte einen Pegel von 55 dB SPL.

Für die Untersuchung harmonischer Entmaskierung wurde die Grundfrequenz des Maskierers prozentual verstimmt, nicht jedoch das Testsignal. Binaurale Entmaskierung wurde durch das Aufprägen einer IPD von 180° auf das Testsignal erreicht.

Die Schwellen wurden einem 3-Intervall 2-AFC 1-up 2-down Verfahren bestimmt. Es nahmen 4 erfahrene Versuchspersonen teil, die jeweils drei Durchläufe absolvierten.

Experiment 1

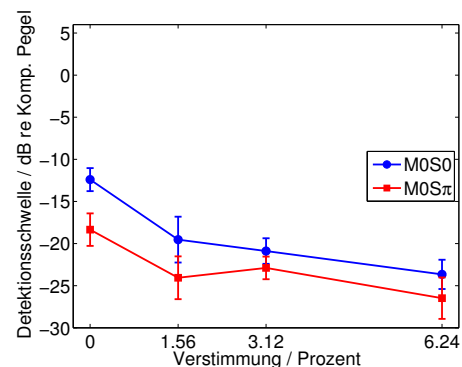


Abbildung 2: Detektionsschwellen für Experiment 1. Schwellen für diotische Konditionen werden durch blaue Kreise angezeigt. Schwellen für Konditionen, in denen die Testkomponente eine interaurale Phasendifferenz von π hatte, werden durch rote Quadrate dargestellt.

Um die Kombination von MMLD und BMLD mit aufgelösten Harmonischen zu untersuchen, wurden $F_0=160$ Hz und $f_t=800$ Hz gewählt. Die Stimuli enthielten also die Harmonischen zwischen 160 und 1440 Hz. Für Verstimmungen von 0%, 1.56%, 3.12% und 6.24% wurden Schwellen gemessen. Die Ergebnisse sind in Abbildung 2 dargestellt.

Die gemessenen Schwellen sind generell niedrig, zwischen -12 und -27 dB relativ zum Komponentenpegel. Mit zunehmender Verstimmung sinkt die Detektionsschwelle. Die harmonische 0%-Kondition zeigt eine Schwelle von -12 dB, die 6.24%-Kondition eine von -24 dB, was einem MMLD von 12 dB entspricht. Die dichotischen Konditionen zeigen generell niedrigere Schwellen als die diotischen

Konditionen. Mit größerer Verstimmung verkleinert sich die BMLD von 6 dB auf 3 dB.

Experiment 2

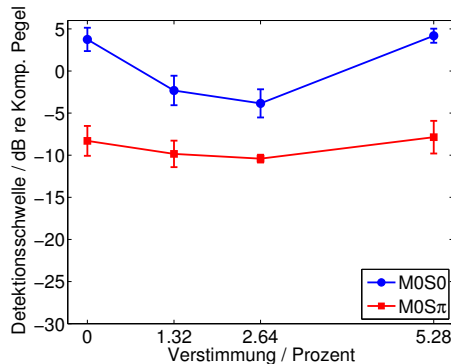


Abbildung 3: Detektionsschwellen für Experiment 2. Darstellung wie in Abb. 2.

In Experiment 2 wurde eine Grundfrequenz von $F_0=40$ Hz verwendet. Die Testkomponente lag wie in Experiment 1 bei 800 Hz, was bei $F_0=40$ Hz der 20. Harmonischen entspricht. Dadurch lagen die Komponenten in einem Frequenzbereich von 640 bis 960 Hz. Die Harmonischen sind durch den Frequenzbereich und die kleinere Grundfrequenz nicht aufgelöst. Es wurden diotische und dichotische Schwellen für Verstimmungen von 0, 1.32, 2.64 und 5.28 Prozent gemessen.

Verglichen mit Experiment 1 wurden insgesamt höhere Schwellen gemessen (s. Abb. 3), die in einem Bereich zwischen 4 und -10 dB liegen. Mit zunehmender Verstimmung sinken die Detektionsschwellen. Bei einer Verstimmung von 5.29% finden sich ähnliche Schwellen wie in der harmonischen 0%-Kondition, was daran liegt dass bei 5.29% Verstimmung die 4. Komponente des Markierers der Testkomponentenfrequenz $f_t=800$ Hz entspricht. Deshalb war für die Versuchspersonen kaum ein Unterschied zwischen Referenzintervall und Testintervall zu hören. Die MOS π Konditionen liefern wie in Exp. 1 tiefere Schwellen als die diotischen Konditionen, zeigen aber kaum Veränderung (max. 2 dB) mit steigender Verstimmung.

Experiment 3

In diesem Experiment wurde wie in Exp. 1 eine Grundfrequenz von $F_0=160$ Hz gewählt, die Frequenz der Testkomponenten war $f_t=4000$ Hz, was der 25. Harmonischen entspricht. Die Komponenten lagen also im Frequenzbereich von 3360 bis 4640 Hz, also in einem Bereich, in dem die Harmonischen aufgrund der Breite der auditorischen Filter unauflöst sind. Für Verstimmungen von 0%, 0.52%, 1.04%, 2.08% und 4.16% wurden diotische und dichotische Schwellen gemessen. Zusätzlich zu den Stimuli wurde ein kontinuierliches Tiefpassrauschen mit einem Pegel von 40 dB SPL dargeboten, um mögliche Ausnutzung von Informationen aus dem tieffrequenten Bereich zu unterdrücken.

Die Ergebnisse sind in Abb. 4 dargestellt. Verglichen

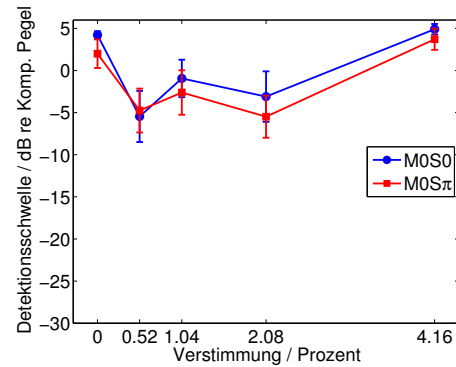


Abbildung 4: Detektionsschwellen für Experiment 3. Darstellung wie in Abb. 2.

mit den anderen Experimenten liegen die Schwellen in einem kleineren Bereich, zwischen -5 und 5 dB. Verstimmung senkt auch hier die Schwellen. Die harmonische und die 4.16%-Kondition zeigen ähnliche Ergebnisse, da die vierte Maskiererkomponente bei dieser Verstimmung auf der Frequenz der Testkomponenten liegt. Im Vergleich zu den anderen Experimenten findet sich eine durchgehend kleinere BMLD von maximal 2.5 dB

Diskussion

Insgesamt zeigen die Ergebnisse, dass sowohl harmonische als auch binaurale Signalmerkmale die Detektionsschwellen senken können. Das Verhältnis zwischen MMLD und BMLD ist komplexer als eine Additivität der Effekte: In Experiment 2 zeigen die dichotischen Konditionen fast keine Abhängigkeit von harmonischen Signaleigenschaften, in Experiment 3 dagegen findet sich wenig BMLD. Dies könnte auf eine parallele Verarbeitung beider Merkmale deuten, nach der schließlich der Effekt, der den größten Gewinn bringt, ausgenutzt wird. Das Ergebnis von Experiment 1 lässt noch keinen endgültigen Schluss zu: hier finden sich in den dichotischen Konditionen noch große MMLDs. Nach Abschluss der noch laufenden Studie sollte die Datenlage nach statistischer Analyse eindeutiger sein.

Danksagung

Diese Studie wurde durch die DFG (SFB/TRR31 "Das aktive auditorische System") unterstützt. Wir danken der Arbeitsgruppe Medizinische Physik und Birger Kollmeier für die Unterstützung und Astrid Klinge und Georg Klump für inspirierende Diskussionen.

Literatur

- [1] A. S. Bregman: Auditory Scene Analysis - The Perceptual Organization of Sound, MIT Press (1994)
- [2] B. C. J. Moore, P. W. Peters, B.R. Glasberg: Thresholds for the detection of inharmonicity in complex tones, JASA 77, 1861-1867 (1985)
- [3] D. M. McFadden: Masking-level differences determined with and without interaural disparities in masker intensity, JASA 44, 212-223 (1968)