

Unerwartete Effekte in der Lautheitswahrnehmung

Jan RENNIES^{1,2}, Jan HOTS³, Jesko L. VERHEY³

¹Boston University, Department of Speech, Language and Hearing Sciences, Boston, MA, USA.

²Fraunhofer IDMT, Project Group Hearing, Speech and Audio Technology, and Cluster of Excellence Hearing4all, Oldenburg, Germany. Email: jan.rennies@idmt.fraunhofer.de

³Otto-von-Guericke University Magdeburg, Department of Experimental Audiology, Magdeburg, Germany. Email: {jan.hots, jesko.verhey}@med.ovgu.de

Einleitung

Lautheit wird seit Jahrzehnten intensiv beforscht und die grundlegenden psychoakustischen Erkenntnisse wurden – vor allem auch durch die Arbeiten der Münchener Schule – in Modelle und Normungsvorschriften zur Lautheitsberechnung umgesetzt. Ein wesentlicher Effekt von grundlegender und praktischer Relevanz (z.B. für die Lärmbewertung) ist die spektrale Lautheitssummation. Diese beschreibt den Effekt, dass sich die Lautheit eines schmalbandigen Signals solange nicht ändert, wie die gesamte Energie in eine Frequenzgruppe fällt. Bei Bandbreiten oberhalb der Frequenzgruppenbreite erhöht sich die Lautheit bei gleichem Gesamtpegel [1].

Spektrale Lautheitssummation kann erfolgreich dadurch modelliert werden, dass das Signal in Frequenzgruppen unterteilt wird, in jeder Frequenzgruppe eine kompressive Transformation von Erregung in sog. spezifische Lautheit (oder Teillautheit) erfolgt und anschließend die spezifische Lautheit über die Frequenzgruppen summiert wird [1]. Dieses grundlegende Prinzip wird heute von allen etablierten Lautheitsmodellen und Standards zur Lautheitsberechnung für stationäre Signale verwendet.

Um über stationäre Lautheit hinausgehende Effekte zu berücksichtigen, wurde der Modellansatz um zeitabhängige Verarbeitungsstufen erweitert, was unter anderem zum Münchener Dynamic Loudness Model (DLM) [2] führte. Der grundsätzliche Ansatz besteht darin, das Prinzip der stationären Lautheitsberechnung in kurzen, über das Signal gleitenden Zeitfenstern anzuwenden. Die daraus resultierende „Instantanlautheit“ wird anschließend mit einer gehörgerechten Zeitkonstante geglättet, um die Kurzzeitlautheit $N(t)$ zu erhalten. In der Regel wird das Maximum oder ein hohes Perzentil (z.B. 95%) der Kurzzeitlautheit verwendet, um die Gesamtlautheit eines Stimulus zu ermitteln. Dies ermöglicht unter anderem die Vorhersage des klassischen Effekts der zeitlichen Lautheitsintegration, d.h., des Anstiegs der Lautheit kurzer Tonpulse mit der Pulsdauer bei konstantem Effektivpegel.

Die beiden Konzepte der spektralen Lautheitssummation und der nachfolgenden zeitlichen Lautheitsintegration sind derart etabliert sowohl in experimentellen Grundlagenwerken als auch in etablierten Modellen, dass sie in der Regel nicht in Frage gestellt werden. In diesem Beitrag werden zwei Effekte näher betrachtet, die mit den klassischen Modellvorstellungen jedoch nicht in Einklang stehen. Diese konkrete Beobachtung und daran anknüpfenden Diskussionen wurden maßgeblich motiviert und inspiriert

durch frühere Beiträge der Münchener Schule. Im Rahmen der Masterarbeit des Erstautors an der Universität Oldenburg, die zeitweise an der TU München durchgeführt wurde, erfolgten dazu erste eigene Untersuchungen und im Anschluss gemeinsame Publikationen mit Münchener Kollegen [3,4]. Fortgesetzt wurden die Arbeiten vorrangig an der Universität Magdeburg durch die weiteren Autoren dieses Beitrags.

Lautheit von Rauschen mit „subkritischer“ Bandbreite

In einem relativ wenig beachteten Beitrag von 1974 stellt Zwicker Lautheitsdaten von frequenzmodulierten Tönen vor [5]. Der Testton bestand aus einem 1,5-kHz Trägerton, dessen Frequenz mit unterschiedlichen Modulationsfrequenzen und einem Frequenzhub von 700 Hz moduliert wurde. Gemessen wurden die Pegel gleicher Lautheit relativ zu einem Bandpassrauschen mit einer Bandbreite einer Frequenzgruppe und einer Mittenfrequenz von 1 kHz. Bereits in der Einleitung beschreibt Zwicker, dass die Ergebnisse im Wesentlichen bereits anhand bestehender Kenntnisse über das Gehör vorhergesagt werden könnten. Bei niedrigen Modulationsfrequenzen nehme das Gehör praktisch einzelne Töne wahr und die Lautheitswahrnehmung entspräche der von Tönen. Bei hohen Modulationsfrequenzen nehme das Gehör effektiv ein Breitbandsignal wahr, dessen Lautheit aufgrund spektraler Lautheitssummation erhöht sei. Nur der bei mittleren Modulationsfrequenzen zu erwartende Übergang zwischen diesen beiden Zuständen sei nicht genau vorhersehbar und müsse daher gemessen werden: „*The reaction of the ear [...] between these two states can not be predicted and has to be measured*“ [5]. Diese gleichsam überzeugte und korrekte Einschätzung drückt in besonderem Maße das Vertrauen in bestehende Modellvorstellungen zum Ausdruck. Tatsächlich zeigen die Daten von Zwicker sehr genau den in der Einleitung vorhergesagten Trend zweier stationärer Lautheitszustände (niedrige Lautheit bei kleinen Modulationsfrequenzen, höhere Lautheit bei hohen Modulationsfrequenzen) mit einem kontinuierlichen Übergang. Spätere Arbeiten sollten zeigen, dass dies auch mit dynamischen Lautheitsmodellen vorhergesagt werden kann [4]. Für diesen Beitrag relevant sind jedoch nicht diese Beobachtungen, sondern vielmehr Datenpunkte, die Zwicker wohl vorrangig als Kontrollbedingungen integriert hatte. Im Rahmen der Messungen verglich Zwicker nämlich zusätzlich die Lautheit von stationären Reintönen bei 1, 1,5 oder 2 kHz und demselben frequenzgruppenbreiten Schmalbandrauschen. Diese Daten zeigen, dass die Töne bei allen gemessenen Referenzpegeln einen niedrigeren Pegel benötigten, um die

gleichen Lautheit wie das Schmalbandrauschen hervorzurufen. Dieser Effekt, der abhängig vom Gesamtpegel in Zwickers Daten einer Pegeldifferenz von bis zu 9 dB entsprach, steht in erheblichem Widerspruch zu den klassischen Modellvorstellungen spektraler Lautheitssummation, nach der Lautheit konstant ist, solange die Energie in dieselbe Frequenzgruppe fällt. In keinem Fall sollte die Lautheit eines Tones höher sein als die eines Rauschens. Interessanterweise wird dieser Effekt in [5] nicht weiter diskutiert. Dies motivierte die Autoren dieses Beitrags dazu eigene Nachforschungen anzustellen, zunächst in Form von Bachelor- und Diplomarbeiten [6,7]. Nachdem erste eigene Daten vorlagen und die Überzeugung hatten reifen lassen, dass dieser Effekt tatsächlich aufzutreten schien, wurde bei einer Gelegenheit in Oldenburg das Gespräch mit Hugo Fastl gesucht, um ihn zu fragen, ob er sich dieser Zwicker-Daten bewusst sei oder gar einen Erklärungsansatz habe. Dies war nicht der Fall, was sich als zusätzliche Motivation erwies systematische Studien anzugehen. Dies Ergebnisse dieser Studien lassen sich in aller Kürze folgendermaßen zusammenfassen [8, 9]:

- die Lautheit von Signalen mit einer Bandbreite unterhalb einer Frequenzgruppenbreite *sinkt* mit steigender Bandbreite bis zur oder sogar über die Frequenzgruppenbreite hinaus;
- diese „mid-bandwidth loudness depression“ ist unabhängig von der Signalreferenz (Ton oder Rauschen), existiert für unterschiedliche Mittenfrequenzen und Referenzpegel und kann sowohl durch eine adaptive Messung von Pegeln gleicher Lautheit als auch durch kategoriale Lautheitsskalierung nachgewiesen werden;
- aktuelle Lautheitsmodelle können „mid-bandwidth loudness depression“ nicht vorhersagen.

Insbesondere der letzte Punkt macht deutlich, dass auch nach jahrzehntelanger Forschung in der Psychoakustik nicht alle Effekte verstanden und vorhersagbar sind.

Lautheit von nicht-synchronen Tonpulsen

Ein weiterer in diesem Beitrag beleuchteter Aspekt geht ebenfalls auf eine frühe Arbeit von Zwicker zurück. Im Gegensatz zum oben beschriebenen Effekt jedoch beschreibt Zwicker in seinem Beitrag von 1969 [10] sehr klar und nachvollziehbar, wie nach seiner Einschätzung die beobachteten Daten zu erklären und effektiv zu modellieren seien. Ziel der Studie „Einfluss der zeitlichen Struktur von Tönen auf die Addition von Teillautheiten“ war die Untersuchung, ob spektrale Lautheitssummation auch dann auftritt, wenn die einzelnen spektralen Komponenten nicht-simultan, d.h., unmittelbar nacheinander, dargeboten werden. Anders als in der Studie zu frequenzmodulierten Signalen macht Zwicker hierzu keine Vorhersage in der Einleitung. Die Daten zeigen eindeutig, dass ein Tonkomplex mit aufeinanderfolgenden Tönen bei gleichem Pegel deutlich lauter ist als ein durchgehender Ton einer festen Frequenz mit derselben Gesamtdauer. Zwickers Erklärungsansatz bestand darin, dass die Trägheit des

Gehörs welche die Grundlage für die zeitliche Lautheitsintegration bildet, bereits auf der Ebene der spezifischen Lautheiten (Teillautheiten) auftreten muss, d.h. vor der Summation der Teillautheiten zur Gesamtlautheit. Nur so lässt sich erklären, dass aufeinander folgende Töne unterschiedlicher Frequenz zu einer Gesamtlautheit führen, die deutlich höher liegt, als anhand der einzelnen Tonpulse zu erwarten wäre.

Eine eigene Studie zu Lautheitsmodellierung dynamischer Signale [3] konnte zeigen, dass Vorhersagen des DLM, welches einen zeitlichen Abklang der Teillautheiten beinhaltet, besser zu Zwickers Daten passen, als Vorhersagen eines konkurrierenden Lautheitsmodells, das zeitliche Lautheitsintegration erst nach der Summation der Teillautheiten modelliert. Nichtsdestotrotz bestand weiterhin eine große Abweichung zwischen den Daten und den Modellvorhersagen, was wiederum eigene experimentelle Studien motivierte, um die zugrunde liegenden Effekte besser zu verstehen. Ausgangspunkt war die Hypothese von Zwicker, dass der Abklang der Teillautheiten die „nicht-simultane spektrale Lautheitssummation“ erklären könnte. Aus ihr folgt direkt, dass der Effekt abnehmen muss, wenn zwischen die Tonpulse eine Pause eingefügt wird und die Pausendauer systematisch erhöht wird. Dies wurde in [11] durchgeführt und bestätigt, wobei als zusätzliche Komplikation bei der Interpretation der Daten auch die Dauerabhängigkeit der spektralen Lautheitssummation [12] berücksichtigt werden muss. Die Studie zeigte ebenfalls, dass aktuelle Lautheitsmodelle nach wie vor nicht alle beobachteten Effekte vorhersagen können.

Zusammenfassung

Die beiden hier beleuchteten Aspekte sind Beispiele dafür, wie die eigenen Forschungsarbeiten der Autoren zur dynamischen Lautheitswahrnehmung maßgeblich durch Studien der Münchener Schule beeinflusst und inspiriert wurden. Teilweise erfolgte dies direkt durch Testen von Zwickers Hypothesen, teilweise jedoch auch durch nähere Betrachtung „seltsamer“ Datenpunkte, die in Zwickers Arbeiten und denen seiner Schüler keine weitere Beachtung gefunden hatten. Aus dieser Neugier und Konstellation ergab sich eine fruchtbare, produktive und hoch geschätzte Zusammenarbeit mit der Münchener Schule, für die die Autoren an dieser Stelle ausdrücklich Hugo Fastl und insbesondere seinem Team von 2008, allen voran Florian Völk, danken möchten. Die Tatsache, dass noch immer nicht alle Aspekte der dynamischen Lautheitswahrnehmung modellhaft beschrieben und damit vorhergesagt werden können, wird auch in Zukunft Motivation und Grundlage psychoakustischer Forschung sein.

Literatur

- [1] Fastl, H. und Zwicker, E. (2007). „Psychoacoustics – Facts and Models,“ 3. Auflage, Springer/Berlin.
- [2] Chalupper, S. und Fastl, H. (2002). „Dynamic loudness model (DLM) for normal and hearing-impaired listeners,“ Acta Acust united Ac 88, 378-386.

- [3] Rennie, J., Verhey, J.L., Chalupper, J., und Fastl, H. (2009). „Modeling temporal effects of spectral loudness summation,“ *Acta Acust united Ac* 95, 1112-1122.
- [4] Rennie, J., Verhey, J.L., und Fastl, H. (2010). „Comparison of loudness models for time-varying sounds,“ *Acta. Acust. Acust.* 96, 383-396.
- [5] Zwicker, E. (1974). „Loudness and excitation patterns of strongly frequency modulated tones,“ in „Sensation and Measurements,“ H.R. Moskowitz et al. (eds), Reidel Publishing Company, Dordrecht-Holland, 325-335.
- [6] Schiller, J. (2009). „Der Einfluss zeitlicher Fluktuationen auf Lautheit,“ Bachelorarbeit, Carl-von-Ossietzky Universität Oldenburg.
- [7] Hots, J. (2010). „Dynamische Aspekte der Lautheitswahrnehmung,“ Diplomarbeit, Carl-von-Ossietzky Universität Oldenburg.
- [8] Hots, J., Rennie, J. und Verhey, J.L. (2013). „Loudness of sounds with a subcritical bandwidth: A challenge to current loudness models?,“ *J. Acoust. Soc. Am.* 134, EL334-EL339.
- [9] Hots, J., Rennie, J. und Verhey, J.L. (2014). „Loudness of subcritical sounds as a function of bandwidth, center frequency, and level,“ *J. Acoust. Soc. Am.* 135, 1313-1320.
- [10] Zwicker, E. (1969). „Einfluss der zeitlichen Struktur von Tönen auf die Summation von Teillautheiten,“ *Acustica* 21, 16-25.
- [11] Heeren, W., Rennie, J., und Verhey, J.L. (2011). „Spectral loudness summation of nonsimultaneous tone pulses,“ *J. Acoust. Soc. Am.* 130, 3905-3915.
- [12] Verhey, J.L., und Kollmeier, B. (2002). „Spectral loudness summation as a function of duration,“ *J. Acoust. Soc. Am.* 111, 1349-1358.