

ワイデックス・Moment™ における パーソナライズされたコンプレッションシステム

2022年10月

ラウラ・ヴィンター・バリング博士、アン・メッテ・ジェッペセン医学博士、イェンス・ブレム・バガー・ニールセン博士

難聴者にとっての基本的な問題は、音を聞き取るために難聴でない人よりも大きな音量が必要なことです。一方、不快な音量となる音のレベル（UCL）は、難聴者とそうでない人とで大きくは変わりません。このため、図1に示すように、難聴者は、小さな音、普通の音、大きな音など、すべての音を聞き分けるためのダイナミックレンジが狭くなっています（赤縦棒）。

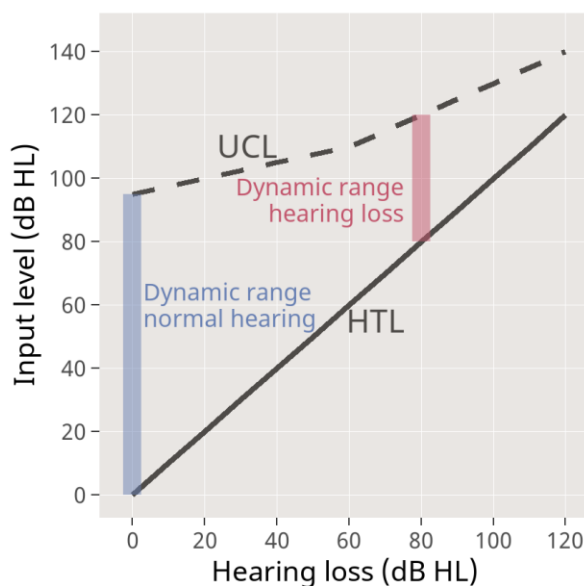


図1：聴覚閾値（HTL、実線）と不快レベル（UCL、破線）の関係から、難聴者（赤）と健聴者（青）のダイナミックレンジの違いを示す。図を簡単にするため、周波数によるHTLの違いは考慮に入れていない。Pascoe (1988)に基づく。

補聴器は、着用者のダイナミックレンジに合わせて全音域を圧縮し、小さな音は比較的大きく、大きな音は小さく、非線形（ノンリニア）に増幅することでこれに対応しています。もちろん、異なる周波数における聴力閾値（HTL）も考慮する必要があります。

今回の WidexPress では、圧縮（コンプレッション）の基礎知識、様々な圧縮の長所と短所、そして Widex の変速コンプレッションの仕組みについて説明します。圧縮はパーソナライズされるべきであるというアイデアは、様々な種類の圧縮に関する幅広い研究結果により支持されています。Widex Moment アプリのアップデートされた AI エンジンによって、圧縮のパーソナライズが実現される革新的な方法を紹介します。

圧縮（コンプレッション）

線形（リニア）増幅と非線形（ノンリニア）増幅のロジックを図 2 に示します。破線は線形増幅を示し、すべての入力レベルに対して同じ量の利得（この例では 15dB）が適用されます。つまり、入力（横軸）が 1dB 増加するごとに、出力（縦軸）が 1dB 増加することになり、入出力比が 1:1 になります。難聴者のダイナミックレンジは限られているため、この方法では、大きな音が大きくなりすぎるか、小さな音が聞き取れなくなるか、あるいはその両方が起こることになります。

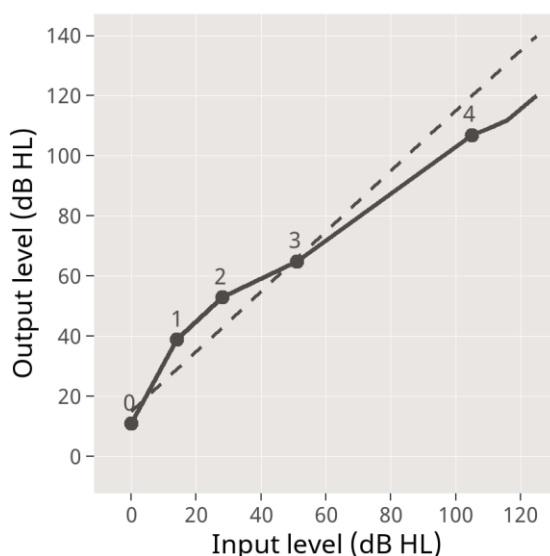


図 2：線形増幅と非線形増幅の比較。非線形のグラフはワイデックス処方式でフィッティングした N3（中程度）難聴の 2kHz の値。実線は、補聴器からの出力と直接透過音の複合効果を示す。

さらに、もっと大きな音は歪んでしまう可能性が高く、実際のケースでは、非常に大きな出力レベルで制限されることとなります。対照的に、図 2 の実線は非線形増幅を示し、異なる入力レベルに対して異なる利得が与えられます。この実線は、N3（中程度）難聴に対してワイデックス処方式で 2kHz の入力を行った場合の入出力曲線です。

このグラフは、入力レベルの範囲が異なると圧縮率も異なることを示しており、圧縮率 1、つまり入力と出力の比が 1:1 であれば、点線で示されるように線形増幅となります。1 以上の比率は圧縮（コンプレッション）を意味し、例えば 2 と 3 の間の実線は、入力が 1.9dB 増加すれば出力が 1dB 増加する、つまり圧縮比 1.9 を表しています。これは、入力が大きくなると出力も大きくなるが、出力の増加は入力の増加より小さいことを意味します。グラフには、圧縮された入出力曲線が、線形増幅の 45 度よりも小さい角度で表示されます。

これに対し、エクспанションは圧縮比が 1 以下の場合であり、入力の増加分よりも出力の増加分が大きいことを意味します。これは、低入力の範囲において、補聴器の内部雑音を含む非常に小さな音の可聴性を減少させる一方で、それらの音よりも大きめではあるがまだ小さな音を聞こえるようにするために使用されます。これは

図 2 の 0 と 1 の間に見られるように、入力が 0.5dB 増加するごとに出力が 1dB 増加する、つまり圧縮比が 0.5 であることを示しています。

番号のついたポイントは、いわゆる圧縮閾値またはニーポイントと呼ばれるもので、圧縮比が変化するポイントです。ワイデックスの補聴器の場合、最も低いニーポイントはかなり低い位置にありますが、それぞれの周波数ごとの聴力レベルによって個別に設定されます。このような細かいカスタマイズがワイデックスサウンドの特長であり、個々のユーザーのためにオーダーメイドしたような素晴らしい音質を実現するための核となっています。

図 2 に示す入出力曲線は、N3 難聴の場合の 2kHz 入力の場合であって、他の難聴程度や周波数の場合は異なります。入出力曲線が 15 バンド全てで連動するように設計されていることと、個々の難聴の程度に合わせて調整されていることが重要です。

圧縮比とニーポイントは、補聴器の音にとって重要な要素ですが、3 つ目の要素として圧縮の速度（速さ）があります。これは、入力レベルの変化に対する補聴器の利得変化の速さを意味します。このような入力レベルの変化には、ドアが閉まる、エンジンが突然止まる、スピーチの後の拍手など、環境音の突発的な大きい変化があれば、会話音声の変調や無音部分など、小さな音の揺らぎもあります。大まかに言えば、ファストコンプレッションはゲインの変化が速く、場合によっては急激に変動する可能性があるということを意味し、スローコンプレッションはゆっくりとしたスムーズな変化を意味します。理想的なのは、入力レベルの変化に対して、可聴性を維持ながら不自然にならず、聞き取りに必要なすべての手がかりを保つことができるコンプレッションシステムです。

圧縮の速度は、アタックタイムとリリースタイムで決定されます。コンプレッサーが入力レベルの増加に対してゲインを下げる時間をアタックタイム、入力レベルの減少に対してゲインを上げる時間をリリースタイムといいます。入力レベルの急激な上昇に対してゲインをどれだけ速く下げるかは、大きすぎる入力から補聴器ユーザーを保護するために極めて重要です。これは、必然的にアタックタイムは速くなければならないことを意味します。そのため、圧縮速度に関する研究や議論の中心は、リリースタイム、つまり入力レベルが低下したときにどの程度の速さでゲインを増加させるかにあります。リリースタイムは最大 20s（秒）で、一般的に短いリリースタイムは 50ms 以下、長いリリースタイムは 500ms 以上と定義されています（Kuk ら、2018 ; Kuk & Hau、2017）。

ファストコンプレッションとスローコンプレッションのメリット/デメリット

ファストとスローのどちらにもメリットとデメリットがあり、圧縮比とニーポイントにも依存しています。システム全体は複雑で多く相互依存しており、様々な組み合わせによる音への影響の研究は明確なものばかりではありません。それでも、いくつか大まかな傾向が浮かび上がってくるので、以下でそれについて説明します。

ファストコンプレッションの第一のメリットは、入力レベルに素早く適応するため、聴取環境の変化をより速く反映させることができることです。例えば、大きな音に続く小さな音をより聞き取りやすくすることができます。また、小さ

な音声（無声子音）を、より大きな音声（母音）より相対的に大きく増幅させることができます。大きな母音と小さな子音との強さの違いを変化させることで、小さな子音をより聞き取りやすくさせます。

しかし、これらメリットには代償があります。利得を高速に変化させると、信号の時間的エンベロープ（包絡線）が変化し、いくつかの点で問題が生じます。第一に、最も一般的に、音質に悪影響を及ぼします。第二に、ワーキングメモリの働きが貧弱なユーザーにとって、劣化した時間的エンベロープは音声明瞭度における問題となりやすくなります。第三に、音声ごとの強度差が変化すると、音声の識別と理解を難しくさせる可能性があります。そして最後に、ファストコンプレッションは静かな環境では周囲の騒音をより認識しやすくし、音声の無音箇所がポンピングノイズとして聞こえてしまう可能性があります。

一方、スローコンプレッションは、高速動作のシステムよりも音声の時間的エンベロープの同一性を保つことができます。これにより、より自然で優れた音質、および聞き心地を向上させることが可能となります。さらに、スローコンプレッションは、音の定位に重要な両耳間の音圧の差をよりよく保ちます。そして、ワーキングメモリの働きが弱く、より重度難聴の補聴器ユーザーにとって、ゆっくり圧縮して音声の時間的エンベロープを維持することは特に有益であると考えられます。

図3は、中程度の背景雑音がある場合の音声信号について、異なるリリースタイムの効果を示したものです。上の波形は、補聴器で録音した信号を、1s（秒）以上の長い（遅い）リリースタイムで、下の波形は、同じ信号を30msの短い（速い）リリースタイムで録音したものです。信号強度の山と谷は、上の長いリリースタイムの方が明らかによく保たれています。一方、下の波形では、ファストコンプレッションによって強度の差が小さくなり、音声の停止時に背景雑音が増幅されてしまっています。スローコンプレッションの方が時間的なエンベロープがよりよく保たれているため、結果として全ての人が良い音質を体験でき、ワーキングメモリの働きが弱い補聴器ユーザーや重度難聴者にとってもよりよい音声の明瞭度を得ることができます。

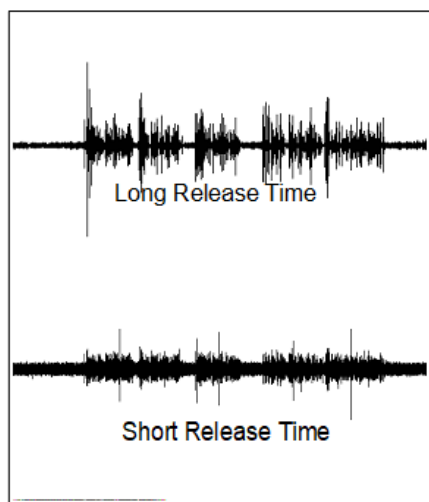


図3: 補聴器で録音した音声信号を圧縮した場合の波形。上は1s（秒）以上の長い（遅い）リリースタイム、下は30msの短い（速い）リリースタイム。時間的なエンベロープ（包絡線）と音声ごとの強度差は、上の長いリリースタイムの方が明らかによく保存されていることが分かる。

スローコンプレッションを単独で使用する場合の主なデメリットは、利得が十分に速く上昇しないため、大きな音の直後の小さな音が聞き取れないことがあることです。さらに、高いワーキングメモリー能力を持つ人の場合、ワーキングメモリー能力が低い人に比べて歪んだ時間的エンベロープの影響を受けにくく、かつ小さな音の可聴性が高まるという恩恵によって、ファストコンプレッションでより良い音声明瞭度を得られる可能性があります。ファストとスローのメリットとデメリットについて図 4 にまとめました。

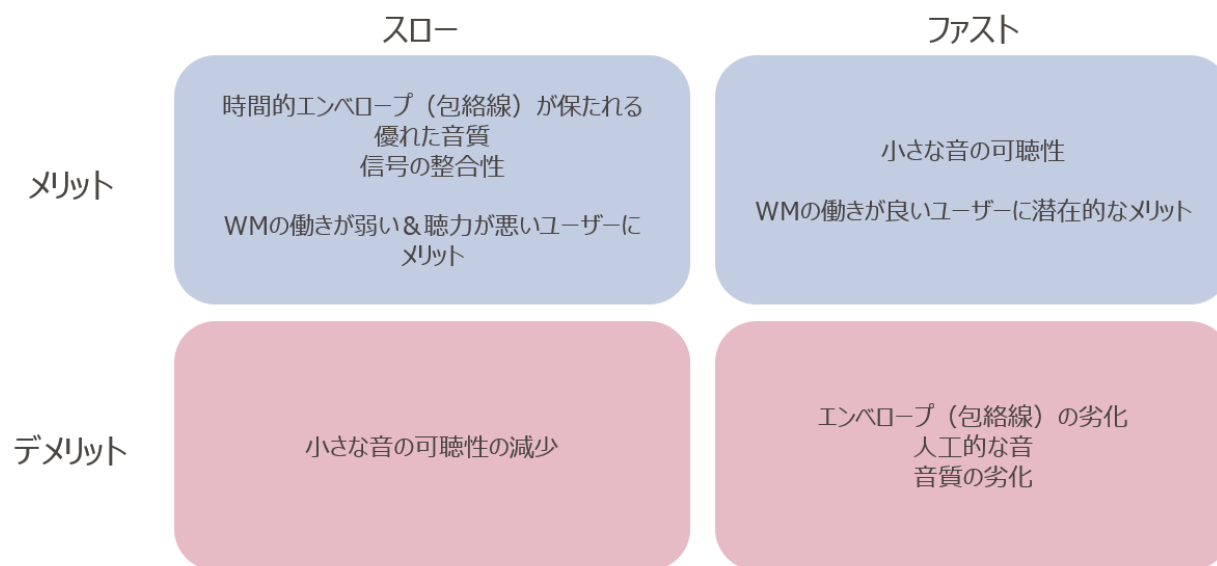


図 4：ファストコンプレッションとスローコンプレッションのメリット・デメリットのまとめ。WM＝ワーキングメモリー。

変速コンプレッション

上記のように、スローコンプレッションとファストコンプレッションのどちらにもメリットがあります。しかし、ファストのデメリットはスローよりも深刻です。ファストコンプレッションは全ての人に対して音質を悪化させ、より重度の難聴者やワーキングメモリー能力が低い人の音声明瞭度を低下させます。

ワイデックスの変速コンプレッション(VSC)が、ファストとスローの両方の要素を持ちながらスローコンプレッションが基本となっているのは、鼓膜に当たる音を最も自然かつ忠実に再現するためです。これはワイデックスが常に目指しているものです。

VSCの基本動作はスローコンプレッションですが、特定の状況下ではより高速な動作を行います。これは、入力レベルの変化とタイミングに依存しており、一般に入力レベルの変化が大きいほど速い変化が許されます。ポイントは、ファストコンプレッションは静かな環境や信号対雑音比（SNR）が良好な環境でのみ有効になることです。これは、SNRの悪い環境で背景雑音が増幅されてしまい、SNRがさらに低下するのを避けるためです。

通常はスローの働きが優位であるため、音質が優先されています。ここにファストの要素が追加されることで、小さい音の可聴性と、特定の状況における音声の明瞭度の向上が期待できます（Kuk ら、2018 ; Kuk & Hau, 2017）。

圧縮のパーソナライズ

このように、変速コンプレッションには確たる後ろ盾がありますが、「ある補聴器ユーザーは速い圧縮を好む/聞き取りが良い」が、「他のユーザーは遅い圧縮を好む/聞き取りが良い」という事実が残っています。スローコンプレッションの音質的なメリットを示すいくつかの研究では、ユーザーのコンプレッションスピードの好みにかんがりのバラツキがあることも示されています。このような好みのバラツキに加え、Souza, Arehart, & Neher（2015）らは、聞き取りの能力差は、ユーザーごとの個人差が大きいことから生じる可能性が高いと主張しています。

個人差に加えて、騒音の種類やさまざまな音声、つまり個人が置かれている環境の違いによっても、好みや聞き取りの能力は変化しやすくなります。つまり、様々なコンプレッションスピードにおける好みと聞き取りの能力は、個人と環境の両方によって変化するため、コンプレッションスピードは様々な環境で個々の補聴器ユーザーのために音をパーソナライズするための理想的なパラメータとなり得るのです。

マイサウンドによるパーソナライズ

ワイデックスの補聴器のパーソナライゼーションは、多くのユーザーと環境に適合するよう自動で音をシームレスに分類、処理し、個々のユーザーに最適なセッティングを提供します。しかし、どんなに優れた自動処理でも、すべてのユーザーのあらゆる状況における聞き取りの意図や好みを予測することはできません。そのため、ワイデックスでは、特定の場面で特別なパラメーターをカスタマイズするオプションも提供しています。

このオプションは、Moment アプリのマイサウンド（図 5）という機能の中にある AI を活用し、働きます。オリジナルのマイサウンドでは、「あなたのためのお勧め」と「あなただけの聞こえを作成」の 2 種類の 방법으로 3 つの周波数帯域のゲインを調整することができます。どちらの場合も、まずユーザー自身が現在の行動・活動と改善意図を示します（図 5a）。「あなたのためのお勧め」を選ぶと、AI アルゴリズムを使って、同じような状況にある他のマイサウンドユーザーの好みに基づいた 2 つの「お勧め」が提示されます。これらビッグデータによる「お勧め」は、ほとんどの場面で役に立ちますが、「あなただけの聞こえを作成」（図 5b の下側）では、より具体的なカスタマイズも可能になっています。「あなただけの聞こえを作成」では、異なる音設定の A-B 比較によって作成が進行します。各ステップにおいて、ユーザーは 2 つの異なる音設定を聞き、画面上のスライダーで好みの度合いを示します（図 5c）。これをもとに、AI エンジンがユーザーの好みに最適化された設定を算出、ユーザーは複雑な操作を意識することなく、その場で音を調整することができます。また、音の体験を言葉に言い換える必要もなく、後日、店頭で補聴器の設定に反映させることも可能です。

これまでのマイサウンドは、低音、中音、高音の周波数帯ごとのゲインを調整することで機能し、音質の評価（図 6）や聞き心地に良い影響を与え、多くのユーザーに聞き取りの改善体験を提供してきました。アップデートされたマイサウンド 2.0（Moment アプリ v1.5）では、3 つの周波数帯のゲイン調整はそのままに、新たにファストコンプレッサーの動作をパラメーターとして追加しました。



図 5：マイサウンドを使ったプログラム作成の流れ。

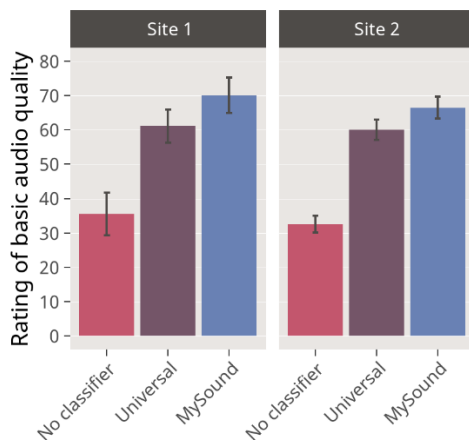


図 6：AI エンジン「マイサウンド」に関する 2 つの研究による音質の評価（サイト 1：n=19、サイト 2：n=20）。

サウンドクラスなし（赤）に比べサウンドクラス付き（紫）は評価が大幅に向上、マイサウンドの AI エンジンで作成された個人用プログラム（青）では、さらに改善が見られた。

マイサウンド 2.0 におけるコンプレッション

マイサウンド 2.0 では、強化された AI エンジンによって、スローコンプレッサーの動作点を軸としてファストコンプレッサーの働きを操作することができるようになりました。全体的なゲインは主としてスローコンプレッサーによって決定されますが、大きな音と小さな音のバランスについて、ユーザーの特定の環境における聞こえ方の好みに合わせて調整することができるようになります。

ファストコンプレッサーの操作が聞こえにどのように反映されるかは、比較的近い音源から来る大きな音と、より遠くから来る小さな音が混在する状況を通じて説明することができます。エクspansionを効かせると入出力曲線の角度が急峻になり（図 2 参照）、相対的に大きな音が強調されます。一方、コンプレッションを効かせると角度が緩やかになり（図 2 参照）、相対的に小さな音が大きく増幅されます。つまり、コンプレッションが効いた音は遠くから聞こえてくる音を聞き取りやすくし、音響的に環境が広がる印象を与えます。逆に、よりエクspansionを効かせると、大きな音を強調することになります。そのため、大きくなりがちな近くの音を強調させて、騒がしい環境での雑音を含む、より遠くの音の干渉を減らすことができます。

重要なのは、オリジナルのマイサウンドと同様、ユーザーが直接信号処理の操作を行わないという点です。実際、有意義に操作できるユーザーはほとんどいません。その代わりに、ユーザーは一連の A-B 比較で提示されるそれぞれの音設定について、どの程度好むかを示していきます。これは比較的簡単な作業ですが、強力な AI エンジンが音設定を作成するために十分な情報を提供します。現在の Moment アプリ v1.5 では、コンプレッションのパラメータは「あなたのためのお勧め」で提供される「お勧め」にはまだ統合されていませんが、「あなただけの聞こえを作成」を使用するユーザーには提供されるよう改良されています。

考察

上記のように、個々の補聴器ユーザーに合わせてコンプレッション機能をカスタマイズすることは、非常に重要です。これまでは、フィッティングソフト上でしかゲインやコンプレッションを調整することができませんでした。つまり、本来調整が必要な環境とはかけ離れた場所で、音の経験を調整に結びつけるという厄介な作業だったのです。マイサウンドの AI エンジンによるファストコンプレッションの調整は、可能なコンプレッション設定の全てをカバーするものではありませんが、全く新しい方法でコンプレッションをパーソナライズすることを可能にする、重要なファーストステップとなるものです。オリジナルのマイサウンドの周波数帯別のゲイン調整の上流に加えられたことによって、マイサウンドによるパーソナライズの可能性を大きく広げることになりました。

さらに言えば、まず店頭で一人ひとりに最適なフィッティングを行った上で、その後ユーザーが自分でその都度パーソナライゼーションを進められるようにしておくこと、これらがワイデックスの補聴器を含む聴覚ソリューション設計の中心的な考え方です。このパーソナライゼーションは、高度な AI、大規模なユーザーコミュニティからのデータ、そして個々のユーザーによるそれぞれの状況における個人的な意見という、業界をリードする組み合わせに支え

られていることが重要です。コンプレッションをパーソナライズ可能なパラメーターのリストに追加することで、このアプローチはさらに重要な一歩を踏み出しました。

参考文献

Balling, L. W., Molgaard, L. L., Townend, O., & Nielsen, J. B. B. (2021). 補聴器ユーザーと人工知能の協働による音の最適化 *Seminars in Hearing*, 42(3), 282-294.

<https://doi.org/10.1055/s-0041-1735135>

Balling, L. W., Townend, O., & Switalski, W. (2019). Widex EVOKE による補聴器の実益ヒアリングレビュー、26(3)、30-36.

Bisgaard, N., Vlaming, M. S. M. G., & Dahlquist, M. (2010). IEC 60118-15 測定手順のための標準的なオーディオグラム *Trends in Amplification*, 14(2), 113-120.

<https://doi.org/10.1177/1084713810379609>

Boike, K. T., & Souza, P. E. (2000). 広ダイナミックレンジ圧縮増幅による音声認識と音声品質評価に対する圧縮比の影響 *音声言語聴覚研究*、43、456-468.

Cox, R. M., & Xu, J. (2010). 短い圧縮解放時間と長い圧縮解放時間 音声理解、実世界での好み、そして認知能力との関連性 <https://doi.org/10.3766/jaaa.21.2.6>

デイヴィス=ベン、E.、ソウザ、P.E.、ブレナン、M.、&ステッカー、G.C. (2009). 重度難聴者の子音認識に及ぼす聴力と多チャンネル広ダイナミックレンジ圧縮の効果.

<https://doi.org/10.1097/AUD.0b013e3181aec5bc>

Dillon, H. (2012). *Hearing Aids* (2nd ed.). Boomerang Press.

Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). 新しい補聴器圧縮解放設定による騒音下での音声認識は、明示的な認知記憶と処理能力を必要とする

<https://doi.org/10.3766/jaaa.18.7.8>

ハンセン、M. マルチチャンネル圧縮時定数が主観的に知覚される音質と音声明瞭度に与える影響 *Killion, Reference Note*, 23, 369- 380.

<https://doi.org/10.1097/01.AUD.0000028009.11739.3E>

Jensen, N. S., Hau, O., Nielsen, J. B. B., Nielsen, T. B., & Legarth, S. V. (2019). 個々のユーザーの好みに基づく機械学習アプローチによる補聴器利得の調整の知覚的効果 聴覚の動向 23、1-23

Jensen, N. S., & Townend, O. (2019). Widex EVOKE における機械学習 : SoundSense Learn の知覚的利点。WidexPress, 42.

Keidser, G., Dillon, H., Dyrland, O., Carter, L., & Hartley, D. (2007). 中重度から高度難聴の補聴器使用者が好む低周波と高周波の圧縮率 <https://doi.org/10.3766/jaaa.18.1.3>

Kuk, F., & Hau, O. (2017). 圧縮速度と認知 すべての人のための可変速コンプレッサー ヒアリング・レビュー、24(3)、40-48.

Kuk, F., Potts, L., Valente, M., Lee, L., & Picirillo, J. (2003). 重度から高度難聴者における順応の証拠 Journal of the American Academy of Audiology, 14(02), 084-099.
<https://doi.org/10.3766/jaaa.14.2.4>

Kuk, F., Slugocki, C., Korhonen, P., Seper, E., & Hau, O. (2018). シングル固定速度コンプレッサーに対するデュアル可変速度コンプレッサーの有効性の評価 Journal of the American Academy of Audiology, 17, 1-17. <https://doi.org/10.3766/jaaa.17.1.27>

Lunner, T., & Sundewall-Thorén, E. (2007). Cognition, Compression, and Listening Conditions の間の相互作用。2 チャンネル補聴器におけるスピーチ・イン・ノイズのパフォーマンスへの影響 Journal of the American Academy of Audiology, 18(07), 604-617.
<https://doi.org/10.3766/jaaa.18.7.7>

ムーア、B. 補聴器における圧縮速度の選択 理論的・実践的考察と個人差の役割 アンプリフィケーションの動向、17(2)、103-112

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1995). 圧縮型補聴器における離床時間の効果 品質に関する一対比較の判断 The Journal of the Acoustical Society of America, 98(6), 3182-3187 <https://doi.org/10.1121/1.413807>

Pascoe, D. P. (1988). 聴覚ダイナミックレンジの臨床的測定と補聴器利得の計算式との関係 第13回 Danavox シンポジウムの議事録

- Souza, P. E., Arehart, K. H., Shen, J., Anderson, M., & Kates, J. M. (2015). 聴覚障害者のワーキングメモリと明瞭度 ワーキングメモリと補聴器で処理された音声の明瞭度 *Frontiers in Psychology*, 6(MAY). <https://doi.org/10.3389/fpsyg.2015.00526>
- Souza, P. E., Arehart, K., & Neher, T. (2015). ワーキングメモリと補聴器処理 文献的知見、将来の方向性、および臨床的応用 *Frontiers in Psychology*, 6(DEC), 1-12. <https://doi.org/10.3389/fpsyg.2015.01894>
- Souza, P. E., Jenstad, L. M., & Folino, R. (2005). 重度聴覚障害者リスナーにおける多チャンネル広ダイナミックレンジ圧縮の使用 音声認識と音質への影響 *Ear and Hearing*, 26(2), 120-131 <https://doi.org/10.1097/00003446-200504000-00002>
- Souza, P. E., & Sirow, L. (2014). 臨床的に適合した補聴器の圧縮パラメータとワーキングメモリーの関連性 *American Journal of Audiology*, 23(4), 394-401 https://doi.org/10.1044/2014_AJA-14-0006
- Townend, O., Nielsen, J. B., & Balslev, D. (2018). SoundSense Learn-リスニングの意図と機械学習 *ヒアリングレビュー*, 25(6), 28-31.