

Uma Breve Revisão da Cognição Auditiva

César R Costa

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Laboratório de Bioinformática e Computação Bio-Inspirada
Núcleo Interdisciplinar de Comunicação Sonora
cesar@nics.unicamp.br

No contexto das ciências cognitivas, o estudo da cognição auditiva busca compreender a percepção sonora e os processos mentais nele envolvidos. De natureza interdisciplinar, esta área de estudos é uma interseção de vários campos de pesquisa que apenas recentemente tem convergido para um núcleo comum. Este artigo apresenta uma breve revisão do assunto com a intenção de promover uma introdução ao leitor, referenciando os trabalhos mais relevantes, considerando o estado de arte atual e perspectivas futuras.

1. Introdução

A capacitação sensorial artificial é de grande interesse para a criação de sistemas autônomos capazes de atuar com a eficiência de organismos vivos naturais em meios reais. O desenvolvimento desta aptidão para os cinco sentidos humanos é uma área de fomento para pesquisa alavancada pelos recentes avanços da robótica, cibernética e computação. O anseio principal é o de dotar sistemas artificiais com a capacidade perceptiva e cognitiva de seres vivos, permitindo-os se comunicar, perceber e compreender o ambiente que os cerca através de estímulos provenientes deste meio.

A questão da cognição auditiva se enquadra neste contexto na preocupação de entender como o ser humano atua sobre a informação sonora e tentar criar modelos computacionais que possam herdar as mesmas capacidades cognitivas. Dada à complexidade deste sentido, ele é estudado em quatro frentes distintas: a música, a psico-acústica, a filosofia da mente e a neurociência.

A música e a cognição musical têm como foco principal de estudo a organização da informação sonora em estruturas de alto nível e sua influência no desenvolvimento cultural e como forma de comunicação. A psico-acústica se preocupa em estabelecer modelos que relacionam características dos estímulos físicos com a forma com que estes são percebidos pelo ser humano. A filosofia da mente nos estudos com a audição trabalha com as representações mentais e o estabelecimento de relações de significado e linguagem utilizando estímulos sonoros. A neurociência tenta estabelecer modelos funcionais inspirados em experimentos práticos com cobaias.

Este artigo busca fazer uma breve revisão introdutória sobre as diversas áreas relacionadas com a cognição sonora. Ele não se caracteriza como uma revisão bibliográfica exaustiva e não cobre de maneira completa todo o estado da arte do assunto, mas atua como um guia introdutório que permite ao leitor se posicionar no assunto obtendo uma visão geral e lhe permite buscar outras fontes para se aprofundar.

O texto está organizado da seguinte forma: na primeira parte as quatro áreas são apresentadas individualmente; na segunda parte é discutido como estas áreas se inter-relacionam; por fim é feita uma breve dissertação sobre o estado de arte atual e as perspectivas futuras.

2. As Quatro Frentes

2.1 Música

A preocupação com a formalização da música data dos tempos dos egípcios antigos. O conjunto de regras e estruturas criadas para organizar o conhecimento sonoro se amadureceu e se estabeleceu de forma tão rígida quanto os cânones matemáticos. Esta é a área de conhecimento relacionada ao estudo da cognição sonora mais estabelecida e de menor dinâmica de desenvolvimento.

A música acusmática surge com paradigmas alternativos ao pregar que a alta organização da música tradicional é imposta por questões culturais. Ela estabelece um nível de organização mais extenso e menos limitado, baseando o arranjo estrutural em uma política mais elementar. Neste sentido, precedido por críticas as limitações do modelo tradicional (Stockhausen, 1957) e (Xenakis, 1971), Scheffer (1966) em seu tratado criou um conjunto taxonômico de objetos musicais. O objetivo é libertar a criação musical de referências ambientais, trabalhando com o conceito de sonoridade em substituição ao de timbre e permitindo-se trabalhar com estruturas de mais baixo nível e fazendo valer do adendo computacional como instrumento generativo. A intenção é fazer música com qualquer tipo de som e não se limitando ao leque de timbres e estruturas da música tradicional. Smalley (1986) formaliza as idéias em uma

taxonomia que ele denomina espectromorfologia.

2.2 Psico-acústica

A psico-acústica é o estudo da percepção subjetiva de sinais acústicos. O objetivo principal é modelar cientificamente os aspectos relevantes para a percepção auditiva humana, incluindo suas limitações e capacidades. Neste sentido, estudos são realizados sobre a própria natureza do fenômeno físico do som e sobre a composição estrutural do aparelho auditivo humano. Dentre os conceitos principais desta área incluem-se o timbre, a intensidade perceptiva (*loudness*) e altura de um determinado som.

Um dos pioneiros na área foi Hermann Von Helmholtz (1877), que criou um modelo simples que associava a composição harmônica do som com a percepção do timbre. Timbre é definido com o conjunto de características físicas que permite associar um determinado som a sua fonte. Pesquisadores que o seguiram, como Rodet (1995) mostraram que muito outros fatores são determinantes na percepção de timbre.

Com relação à percepção de altura e intensidade, os estudos de Fletcher e Munson (1933) mostraram que a intensidade percebida é dependente da altura do som. Esta dependência foi quantificada nas curvas de *loudness*, atualmente regularizadas pela ISO como ISO226:2003. Devido ao atrelamento das duas qualidades, foi definida a unidade *phon* como uma medida de intensidade subjetiva equivalente à intensidade percebida por um som de 1dB e 1kHz.

Outro fenômeno perceptivo verificado em relação à altura e à intensidade é o mascaramento. Quando dois sons são

apresentados simultaneamente, uma diferença acentuada de intensidade pode fazer com que apenas o mais intenso seja percebido. Este fenômeno é fortemente influenciado pelo espectro de frequência dos dois sons, e em sons de espectro rico existe a ocorrência de mascaramento entre as próprias parciais. Por este motivo, a pura análise por transformadas como a de Fourier ou Wavelets não expressam exatamente como as amostras são percebidas. Modelagens psico-acústicas podem ser utilizadas para criar representações mais reais da percepção sonora.

Derivada da mesma questão do mascaramento, os estudos de Plump (1976) estabeleceram os fatores que fazem tons diferentes sejam percebidos de forma dissonante ou consoante. Ele verificou a existência de uma banda crítica de frequência em que dois sons simultâneos causam uma sensação desagradável. De seus estudos foi possível explicar muitas das relações de harmonia da música tradicional.

O nível de conhecimento referente à psico-acústica já se amadureceu a ponto de gerar modelos eficientes aplicáveis em situações práticas. Um exemplo são os codecs de áudio digital como MP3 ou Ogg Vorbis que utilizam transformadas e filtragens em conjunto com o conhecimento do modelo auditivo para eliminar informação de pouca relevância para a percepção sonora e obter altas taxas de compressão. O conhecimento do comportamento da percepção auditiva abriu caminho para o desenvolvimento de arquiteturas de auditórios e ambientes públicos acusticamente mais eficientes e agradáveis.

2.3 Filosofia da Mente

A cognição auditiva não foi uma preocupação nos primeiros modelos. Nos

modelos cartesianos iniciais a questão auditiva estava muito além do que se trabalhava. A inserção lingüística se restringia a questão verbal e a criação de símbolos seja visual ou sonoro não era considerada. Nesta época, a única interseção se dava a partir dos estudos de cognição musical em que a música era utilizada como uma ponte entre o som e os modelos de informação. A associação fonética com a língua falada, por mais que algumas vezes fosse considerada, necessitava de um ferramental que levasse ao domínio verbal.

A questão auditiva só se tornou prática com o adendo da semiótica que criou a possibilidade de sons e imagens atuarem como signos propensos ao processamento mental. Braun (1991) fez um estudo sobre a atuação cognitiva dos sentidos. No modelo modular de mente de Jackendoff (1993) a música foi considerada como uma linguagem da mente. Neste contexto é possível estender o conceito de música para objetos sonoros genéricos organizados. Santaella (2001) descreve a teoria das matrizes da linguagem e do pensamento em que inclui a linguagem sonora, visual e verbal. Ela apresenta a linguagem sonora como a mais primordial com forte vocação sintática, mas menor vocação referencial ou de discurso que a visual ou verbal.

2.4 Neurociência

A neurociência tenta mimetizar as funcionalidades do sistema auditivo através da observação do comportamento neural. Arquiteturas conexionistas são desenvolvidas com a principal finalidade de se obter a funcionalidade desejada e possibilitar o estudo a partir do modelo criado.

Um exemplo é a cóclea artificial desenvolvida por Kern et al. (2002)

cujos estudos levaram a avanços no entendimento do funcionamento da cóclea natural (Kern et al. 2004).

Existem trabalhos que utilizam estratégias conexionistas e auto-organização na obtenção de técnicas criativas de síntese sonora como em (Costa et al. 2005).

3. A Convergência

Apesar da independência de origem e de desenvolvimento das quatro áreas, têm surgido questões que transitam entre elas.

A música tem se aproximado da psico-acústica com o surgimento da música acusmática. O estabelecimento de uma organização de nível mais baixo a aproxima do sinal sonoro, e da psico-acústica a música extrai conhecimentos para desenvolver estas novas estruturas.

A filosofia da mente se comunica com a música na tentativa de compreender os processos cognitivos envolvidos na organização musical. O estudo da chamada cognição musical tenta compreender os processos cognitivos envolvidos na organização musical. Com relação aos estudos da música tradicional, os trabalhos se seguem no sentido de identificar como as estruturas musicais são interpretadas pelo ser humano, os processos envolvidos na composição e na leitura. Downling (2001) faz uma boa revisão histórica desta convergência com foco principal nas questões da psicologia. Rowe (1993) cria uma série de sistemas musicais interativos aos quais são atribuídas funções de escuta e composição a partir de modelos baseados no comportamento humano. Sloboda (1988) estuda estes processos generativos sob uma perspectiva da psicologia.

Neste contexto, a formação do conceito de gêneros musicais é uma evidência do processamento em alto nível realizado pelo ser humano. Existem sistemas automáticos de composição que tentam extrair características que possam caracterizar um tipo de gênero musical a partir de um conjunto de músicas amostrais.

Quanto à criação de modelos, Desain et al. (1998) faz uma boa discussão e tenta estabelecer os problemas principais e alguns métodos a serem seguidos na obtenção de modelos práticos. Não se estabelece algo prático, mas serve como boa base de suporte teórico.

A filosofia da mente também conversa com a psico-acústica na busca da criação de modelos interpretativos do fenômeno sonoro e da mimetização de habilidades humanas dependentes da audição.

Uma área diretamente ligada à psico-acústica, mas que busca uma maior integração com os estudos cognitivos é a capacidade de análise de cena acústica (*Auditory Scene Analysis*). A capacidade de análise de cena é uma faculdade do ser humano de utilizar seus sistemas sensoriais para criar uma representação do ambiente que o cerca.

Bregman (1990) objetiva a criação de uma teoria única que seja capaz de sintetizar a psico-acústica com a cognição acústica em uma única teoria de audição. Ele é parcialmente bem sucedido uma vez que consegue organizar os conhecimentos da psico-acústica e definir conjuntos de capacidades que se baseiam nos princípios conhecidos. Dentre elas, destaca-se o poder do ser humano de segregar objetos sonoros e estabelecer identidades contínuas. Capacidade análoga no campo visual à identificação

de objetos como formas geométricas ou carros.

Contudo, não é estabelecido um modelo representacional do sistema auditivo. Mas ele deixa o terreno preparado para a integração das ciências cognitivas com o conhecimento da psico-acústica.

Das quatro áreas, a neurociência se apresenta como a mais imatura e isolada. A imaturidade está relacionada com o fato dos estudos iniciais serem muito recentes. O isolamento é uma consequência da própria natureza da neurociência que tenta estabelecer modelos a partir de observação estrutural, de certa forma ignorando as outras áreas. Por isso a neurociência acaba atuando principalmente como um suporte de desenvolvimento ao criar conhecimento próprio que pode ser estudado pelos outros campos.

4. Conclusão

O impulso original das quatro áreas foi diferente, mas a essência de estudar a relação humana com o fenômeno sonoro é comum. Destas, a música se apresenta como a mais madura dentro de suas propostas com o estabelecimento de todo um ferramental teórico que permite compreender e sustentar a criação. A psico-acústica também atingiu um alto grau de maturidade o que tem pressionado seus estudos a buscar novos horizontes nas ciências cognitivas. Os modelos computacionais psico-acústicos já são uma realidade.

A filosofia da mente e a neurociência ainda estão a um passo atrás. A filosofia de mente devido ao fato que a efervescência de novas teorias ainda é muito elevada (apesar de já ter uma forte carga teórica no campo da psicologia). Os modelos computacionais não tratam o som em sua característica sintática,

mas o transformam em estruturas musicais ou verbais. A aplicação de modelos semióticos que sejam capazes de considerar mídias como signos estabelecidos poderão levar a teoria cognitivista a gerar modelos computacionais plenos para a aplicação auditiva.

A neurociência já foi capaz de criar modelos funcionais para alguns processos auditivos, mas apenas para funções mais isoladas. Quando estes se estenderem para funções mais complexas estes poderão ser integrados aos modelos cognitivistas.

5. Bibliografia

Braun, E. (1991) *El saber y los sentidos*. México: Fondo de Cultura Económica.

Bregman, A. S., *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press, 1990.

Bregman, A. S. and Dannenbring, G. 1973. The effect of continuity on auditory stream segregation. *Perception Psychophysics* 13: 308-12.

Costa, C. R., Caetano, M., Manzolli, J., and Von Zuben, F. J. Self-Organizing Topological Timbral Design Methodology Using a Kohonen Neural Network. Em *Proceedings of the 10th Brazilian Symposium on Computer Music (SBCM)*, Belo Horizonte, Brazil, 2005c, 94-105.

Dowling, W. J. (2001). Music perception. In E. B. Goldstein (Ed.) *Handbook of perception*. Oxford : Blackwell, pp. 469-498.

Desain, P.W.M., Honing, H.J., Vanthienen, H.A., & Windsor, W.L. (1998). Computational modeling of music cognition: Problem or solution?. *Music Perception*, 16 (1), 151-166.

- Fletcher, H., Munson, W. A. (1933) "Loudness, Its Definition, Measurement and Calculation" *The Journal of the Acoustical Society of America* -- October 1933 -- Volume 5, Issue 2, pp. 82-108
- Huron, D. "Albert S. Bregman: Auditory Scene Analysis: The Perceptual Organization of Sound [review of]." *Psychology of Music*, Vol. 19, No. 1, pp. 77-82.
- Jackendoff, R. (1987) *Consciousness and the computational mind*. MIT Press.
- Kern, A. and van der Vyver, J.-J. and Stoop, R. (2002) "Towards A Biomorph Silicon Hopf Cochlea" , *NDES IEEE Conference on Nonlinear Dynamics of Electronic Systems*.
- Kern, A. and Stoop, R. (2004) "Nonlinear Dynamics of Cochlear Information Processing" , *Proceedings of the IEEE conference NDES 2004*, Evora, 2004 190-193.
- McAdams, S., Bigand, E. (1993) *Thinking in Sound: The Cognitive Psychology of Human Audition*, Oxford University Press, Oxford
- Plomp, R. (1976) *Aspects of tone sensation: A psychophysical study*, Academic Press.
- Rodet, X., Ph. Depalle, Garcia, G. (1995) *New Possibilities in Sound Analysis and Synthesis*, ISMA 1995
- Rowe, R. (1993) *Interactive Music Systems: Machine Listening and Composing*, Cambridge, MA: The MIT Press.
- Santaella, L. (2001) *Matrizes da Linguagem e do Pensamento*. Editora Iluminuras.
- Schaeffer, P. (1966) *Traité des Objets Musicaux*. Editions Du Seuil.
- Sloboda, J.A. (1988). *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*. Oxford, Oxford University Press.
- Smalley, D. (1986) Spectromorphology and Structuring Processes. In *The Language of Electroacoustic Music*, Org. Emmerson S., Harwood Academic Publishers, New York.
- Stockhausen, K. (1957) "...how time passes...", *Die Reihe #3*, English trans. Cornelius Cardew 1959, p. 10
- Von Helmholtz, H. (1877) *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Dover Publications. New York. 1954 reprint.
- Xenakis, I. (1971), *Formalized Music*, English trans., Bloomington: Indiana University Press.