

# Linguística, música e cognição humana: da representação estática à interação dinâmica

---

DOI: <http://dx.doi.org/10.21165/el.v48i1.2388>

**Verônica Penteado Siqueira<sup>1</sup>**

## **Resumo**

O objetivo deste trabalho é expor algumas das principais abordagens da ciência da cognição sobre as línguas naturais e a música, a fim de traçar um elo entre linguística, música e cognição humana, e, com isso, traçar o percurso realizado desde a pesquisa de Iniciação Científica até a seleção do objeto de estudo da pesquisa de mestrado. As abordagens serão revistas com o propósito de explorar tal objeto: a fala sincronizada (e a *joint speech*), vista a partir do paradigma dinâmico da cognição e da linguagem. Dessa maneira, procura-se defender a relevância da abordagem dinâmica para tratar da relação entre a música e a linguagem, bem como da própria natureza da fala, e apresentar a fala conjunta (*joint speech*) como um objeto de estudo empírico para investigar questões que vão além da fala e da linguagem.

**Palavras-chaves:** cognição musical; cognição; fala; dinamicismo.

---

<sup>1</sup> Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; [veronica.siqueira@usp.br](mailto:veronica.siqueira@usp.br); <https://orcid.org/0000-0001-8793-7936>

## Linguistics, music and human cognition: from a static representation to a dynamical interaction

### Abstract

The aim of this paper is to present the main approaches of the cognitive sciences about natural languages and music, in order to outline a link between language, music and human cognition, and, with that, sketch the trajectory made from the Iniciação Científica research to the selection of the Master's object of study. The approaches will be reviewed with the goal of exploring such object of study: the synchronized speech (and the joint speech), analyzed from a dynamical perspective of cognition and language. Thus, we support the relevance of the dynamical approach to account for the relationship between music and language, as well for the nature of speech itself, and introduce the joint speech as a productive object of empirical study to investigate matters that go beyond speech and language.

**Keywords:** musical cognition; cognition; speech; dynamical approach.

### Introdução

Há uma intercambialidade, ou até mesmo uma confusão, entre os termos linguagem, língua e fala, de um lado, e música, canção e canto, de outro. Não se almeja esclarecê-los exaustivamente aqui, mas o tema do trabalho pede ao menos uma breve (tentativa de) problematização do assunto, antes que possamos adentrar nele de fato.

No senso comum, os termos linguagem e fala são muitas vezes intercambiáveis. Em inglês, o termo para linguagem é ambíguo; podemos empregar o termo *language* referindo-se tanto à capacidade linguística quanto ao idioma com o qual cada pessoa comumente se comunica, seja o português, o coreano, o swahili ou o *American Sign Language*. Em português, essa ambiguidade desaparece; linguagem se refere à capacidade, ao passo que língua se refere, por exemplo, ao português. Quanto à fala, Albano (1990, p. 20) traz uma visão interessante, afirmando que “se chega à linguagem tocando a fala de ouvido”; enquanto a fala diz respeito aos movimentos sensorio-motores, a linguagem refere-se ao sistema simbólico que emerge a partir desses movimentos. Com isso, Albano (1990) defende a tese de que a fala inicial do falante – as vocalizações e a audição (e a gesticulação e a visão, no caso das línguas de sinais) – é parte do desenvolvimento da linguagem, uma vez que a abstração desses padrões concretos leva ao sistema simbólico que é a linguagem.

Quando se trata dos termos música e canção, aí é o português que assume a ambiguidade. Se em inglês a distinção entre *music* e *song* é clara, referindo-se o primeiro aos sons produzidos por instrumentos musicais ou pela voz, ou pela combinação desses, e o

segundo, a uma peça musical composta por palavras, os vocábulos música e canção se misturam, em especial, no português brasileiro. Dada tamanha difusão e popularidade da canção no Brasil, os termos são usados alternadamente quando nos referimos à canção, apesar de ela ser, na verdade, hipônimo de música. No que concerne à relação entre fala e canto, em geral, ou ambos os fenômenos são considerados pertencentes a domínios diferentes ou a aproximação entre eles é superficial; ambas as abordagens são insuficientes para tratar da relação entre essas duas realizações sonoras (RAPOSO DE MEDEIROS, 2014).

Para além da complexidade com a qual nos deparamos ao definir esses termos, música e linguagem são fenômenos que suscitam questões não só a respeito das suas particularidades, mas também a respeito do funcionamento do cérebro e da cognição humana. É possível argumentar que música e linguagem são universais, em diversos sentidos; ambos são sistemas sonoros, que todos os seres humanos possuem, formados por unidades sonoras discretas, e percebidos, organizados e interpretados pela cognição (PATEL, 2003; PATEL, 2008; FEDORENKO *et al.*, 2009; JACKENDOFF, 2009). Pode-se dizer que são universais na medida em que são encontradas em diversas culturas do mundo, ao longo da história da humanidade. McDermott e Hauser (2005) argumentam a favor de tal universalidade para a música, focando-se sobre explicações evolutivas da origem e do desenvolvimento da música. Também pode-se argumentar que é universal a sua diversidade, isto é, que é possível detectar diferenças entre as características das diversas músicas e línguas naturais do mundo<sup>2</sup>.

Por outro lado, é possível também afirmar que a música é diferente da linguagem, baseando-se no argumento de que todos os seres humanos são seres linguísticos, mas nem todos têm habilidades musicais. No entanto, é enganoso pensar que “ser musical” está relacionado estritamente a cantar bem ou tocar um instrumento musical. Os seres humanos são seres musicais no sentido de que são capazes de compreender e apreciar música, assim como produzi-la, em diferentes níveis de habilidade. Confundir habilidades profissionais, adquiridas por meio de treinamento formal, com a capacidade natural de ouvir e apreciar música, bem como de acompanhá-la, seja por meio do canto ou da dança, é um equívoco. A não ser que o indivíduo possua alguma deficiência musical, todos os seres humanos parecem ser igualmente inclinados à música (PERETZ, 2006). Em experimentos comparando as habilidades de músicos e não-músicos, observa-se que ambos apresentam habilidades semelhantes em tarefas de produção e percepção (BIGAND; POULIN-CHARRONAT, 2006). A capacidade de ouvir e produzir música está relacionada a conhecimentos e processos mais centrais, presentes em todos os seres humanos que não sofrem de nenhuma deficiência musical. A habilidade de músicos profissionais ou amadores, em comparação a de não-músicos, não provém desse sistema central de conhecimentos musicais, mas sim do ensino e treinamento formal em música.

---

2 Ver Nettl (1983) para uma visão da etnomusicologia sobre a universalidade e diversidade da música.

Discussões a respeito da natureza inata da música e de sua origem, em comparação à capacidade linguística, à parte (MCDERMOTT, HAUSER, 2005; PERETZ, 2006), a relação entre música e linguagem pode ser abordada por diferentes pontos de vista. Do ponto de vista estrutural, nem todos os paralelismos são possíveis, mas a partir da perspectiva do som, essa comparação se torna mais visível, uma vez que a substância é a mesma. Obviamente, traçar paralelos entre o fonema e a nota musical, por exemplo, deixa de ser produtivo ou mesmo possível, sem fazer comparações demasiadamente forçadas. A começar pelas descobertas de Sundberg (1987), estudos mostram como canto e fala se adaptam uma à outra, respeitando suas exigências articulatórias e prosódicas (RAPOSO DE MEDEIROS, 2002; KOLINSKY *et al.*, 2009; RAPOSO DE MEDEIROS; CABRAL, 2018). Entre teses e dissertações, contam-se até o momento 27 trabalhos que investigam o canto, comparando-o à fala ou analisando seus aspectos acústicos ou fonológicos<sup>3</sup>.

Em termos abstratos, nem sempre há comparação. Muitos autores comparam as duas capacidades em termos da sua sintaxe e do seu processamento, como Lerdahl e Jackendoff (1983) e Patel (2008), cujas propostas serão abordadas nas próximas seções. É possível também observar tal relação em termos mais estreitos cognitivamente, como atenção, memória e recursividade. Indagações a respeito do tipo de significado transmitido tanto pela música quanto pela linguagem também percorrem o estudo da relação entre essas capacidades e, nesse ponto, ambas não parecem ter muito em comum.

Tendo essa discussão em vista, este trabalho tem como objetivo apresentar algumas abordagens que tratam da relação entre a cognição musical e a cognição linguística, e, com isso, traçar o percurso realizado pela autora desde a pesquisa de Iniciação Científica até a elaboração do tema da pesquisa de mestrado<sup>4</sup>. Durante a IC, foi feito um levantamento bibliográfico de diferentes abordagens das Ciências da Cognição que tratassem da relação entre música e linguagem, a fim de traçar um elo entre as duas capacidades. A pesquisa de mestrado não mais concerne aos estudos musicais, mas esses serviram como ponte, assim como os estudos sobre dinamicismo serviram como base, para pensar questões da fala. Neste artigo, as abordagens serão revistas com o propósito de explorar o objeto de estudo da pesquisa de mestrado – a fala sincronizada (e a *joint speech*) –, vista a partir do paradigma dinâmico da cognição e da linguagem. Dessa maneira, procura-se defender a relevância da abordagem dinâmica para tratar da relação entre a música e a linguagem, bem como da própria natureza da fala, e apresentar a fala conjunta (*joint speech*) como um objeto de estudo empírico para investigar questões que vão além da fala e da linguagem, permitindo abordar discussões tanto sobre a cognição humana quanto sobre a organização social.

---

3 Informação fornecida por Raposo de Medeiros em comunicação pessoal, em agosto de 2018.

4 A pesquisa de mestrado é financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Abordagens teóricas

No recorte teórico feito na pesquisa de IC, é possível, de maneira geral, distinguir as abordagens entre aquelas que são mentalistas e aquelas que, diferentemente, abordam a cognição humana segundo o paradigma dinâmico. Ambas querem explicar um comportamento inteligente, como o linguístico e o musical, mas o fazem de maneiras diferentes.

A visão tradicional da cognição, com origens no trabalho de Fodor (1983) sobre a modularidade da mente, trata da descrição da mente baseada na metáfora computacional, isto é, na ideia de que a mente opera como um computador. Nessa abordagem, lidamos com um *input* simbólico, que é interpretado, processado e armazenado e, assim, determina um *output* da ordem do físico, da realização. Essa é, por exemplo, a diferença tradicional entre a Fonética – da ordem do físico – e a Fonologia – da ordem do simbólico e do abstrato.

O dinamicismo, por outro lado, mostra-se como uma alternativa ao cognitivismo, ao colocar em questão algumas de suas ideias, como a representação mental, o internalismo, o reducionismo mente-cérebro e o computacionalismo (para algumas referências sobre o cognitivismo, ver FODOR, 1983; MILLER, 1951; CHOMSKY, 1957; NEWELL; SIMON, 1972). Segundo a perspectiva dinâmica da cognição, tanto a linguagem quanto a música são vistas como capacidades emergentes do funcionamento cognitivo geral e da interação social, e não somente como capacidades emergentes de um módulo mental, como tradicionalmente visto. Esse paradigma traz uma visão diferente do cognitivismo aos estudos cognitivos da música e da linguagem ao distanciar-se dessa visão psicológica e representacionista, cuja necessidade ou utilidade para a explicação das capacidades pode ser questionada (HASELAGER, 2004). Nessa perspectiva, o estudo da relação entre música e linguagem demanda uma visão coordenada dessas atividades, que leve em conta tanto a ação dos falantes quanto dos ouvintes na composição do processo como um todo.

Enquanto uma abordagem mentalista toma a representação como uma intermediária entre a percepção e a ação, especificando informações a respeito do mundo de forma independente do contexto e da ação, uma abordagem dinâmica dispensa esse conceito, preferindo explicar os fenômenos segundo seu contexto e sua ação. Questiona-se aqui se não é possível superar a noção de representação mental, isto é, se não nos tornamos muito dependentes desse conceito para explicar os fenômenos não só linguísticos, mas também da cognição humana.

## Teoria Gerativa da Música Tonal (GTTM)

Tomou-se como ponto de partida para leitura e análise da relação entre a música e a linguagem a Teoria Gerativa da Música Tonal (GTTM, *Generative Theory of Tonal Music*, em inglês), divulgada em 1983, e formulada pelo linguista Ray Jackendoff e pelo músico Fred Lerdahl. Foi a primeira teoria musical que se propôs a formular uma descrição das estruturas mentais correspondentes às músicas ouvidas por um indivíduo. Esse aspecto mental e formalista da teoria não é desmotivado; como o próprio nome sugere, ela parte da teoria gerativista da linguagem elaborada por Noam Chomsky (ver, por exemplo, CHOMSKY, 1965, 1968, 1975), que também procura descrever e explicar as estruturas mentais responsáveis pela composição e compreensão de uma sentença linguística. Ambas as teorias têm em comum, portanto, o mentalismo e o formalismo; possuem a mesma pretensão de formular uma gramática capaz de explicar as capacidades linguística e musical, e assim o fazem estabelecendo uma relação entre a substância propriamente dita e a representação mental desta.

Observa-se, então, que, para os autores, a capacidade musical é tão abstrata quanto a capacidade linguística. Ademais, música e linguagem funcionam de modos similares no que diz respeito à sua organização e ao modo como as suas partes se articulam para formar um todo. Assim como a linguagem é resultado do processamento simultâneo de estruturas independentes e hierárquicas – fonologia, sintaxe e semântica – as estruturas de grupo, metro e *pitch* se organizam para formar a música (JACKENDOFF; LERDAHL, 2006).

No entanto, quando nos afastamos das características mais gerais para analisar as especificidades de cada faculdade, e abordamos as características estruturais e funcionais, diminuem-se os pontos comuns entre música e linguagem. Segundo os autores, essas diferenças estruturais são resultado de uma divergência de função fundamental: enquanto a linguagem expressa pensamentos proposicionais (sendo proposição não uma sentença linguística, mas sim um tipo de estrutura conceitual mais abstrata), a música acentua as emoções, ou como os autores preferem chamar, afetos. Isto é, enquanto a linguagem é a combinação entre sons e conceitos, a música combina sons a expressões afetivas.

Tal diferença dita “fundamental” por Jackendoff e Lerdahl (2006) coloca a componente sintática também como um ponto de distinção. Apesar dessa diferença, Jackendoff e Lerdahl (2006) reconhecem a semelhança entre a estrutura linguística e a musical na medida em que ambas são hierárquicas e recursivas. Propõem, então, a chamada estrutura de prolongamento (*prolongational structure*), característica mais próxima na música à sintaxe linguística.

Essa estrutura se assemelha muito às árvores sintáticas do gerativismo, organizando as relações entre as notas de maneira hierárquica e recursiva. No entanto, as semelhanças se restringem a isso. Não há algo nesta estrutura musical que a assemelhe à Teoria X-barras da linguística gerativa<sup>5</sup>; por exemplo, o acorde de dó maior não se organiza tendo a nota dó como núcleo, ao contrário, por exemplo, de um sintagma nominal, que tem como núcleo um nome. Assim, a teoria restringe-se aos componentes hierárquicos da música, sobre os quais são aplicadas regras. Isto é, a superfície musical<sup>6</sup> é o *input*, enquanto a estrutura ouvida pelos ouvintes é o *output*.

A principal regra do sistema – e o aspecto fundamental que diferencia essa teoria musical da teoria linguística de Chomsky – são as regras de preferência (*preference rules*), que não possuem correspondente na teoria linguística. Essas regras se referem aos padrões de tensão e relaxamento da música, isto é, às estabilidades e instabilidades entre as notas musicais. Segundo a GTTM, as relações entre as alturas musicais e sua posição na hierarquia dependem da sua estabilidade na estrutura. É feita uma distinção entre notas mais estruturais, mais importantes para a formação da melodia e de notas mais “ornamentais” para a melodia (o que seria semelhante a dizer que um certo componente causaria um problema maior na estruturação de uma sentença, como um verbo, em comparação a um adjetivo). Quanto mais estável o fenômeno, mais alta é a sua posição na estrutura, expressando um ponto de relaxamento no contorno musical. Inversamente, quanto mais instável a nota, mais baixa é a sua posição na hierarquia, expressando pontos de maior tensão<sup>7</sup>.

## Teoria do Espaço Tonal

A estrutura de prolongamento da GTTM e as relações hierárquicas entre as alturas musicais expressas por ela suscitam questões de natureza empírica, e a Teoria do Espaço Tonal (TPS, *Tonal Pitch Space*, em inglês), formulada por Lerdahl (1988, 2001), procura justamente resolver algumas das idealizações da teoria musical gerativa por meio da análise quantitativa das relações tonais da música.

Tomando um centro tonal como ponto de referência, também chamado de tônica, é possível medir a distância cognitiva entre os componentes, a qual é responsável pela

---

5 A Teoria X-Barra trata da organização de categorias sintáticas em uma estrutura hierárquica. Nessa estrutura, cada sintagma (XP) possui um núcleo (X), que determina as relações com outros sintagmas.

6 Na GTTM, entende-se por superfície musical a sequência de eventos musicais, acima do nível psicoacústico, sobre a qual atua tal sistema de regras, não devendo ser confundida com o conceito de estrutura de superfície da linguística gerativa.

7 Para mais detalhes sobre as relações hierárquicas da estrutura de prolongamento, sugiro o capítulo 5 de Jackendoff e Lerdahl (1983).

estabilidade do evento e, portanto, pelos padrões de tensão e relaxamento da música. Uma altura (*pitch*) não é ouvida apenas em relação ao centro tonal (altura fixa), mas também em relação às alturas mais estáveis presentes no espaço tonal. Assim, cria-se uma expectativa em relação às notas seguintes, de acordo com a “atração” que uma nota causa sobre a outra.

Observa-se que Lerdahl (2001) entende a capacidade musical como produto de dois aspectos, acústico e cognitivo. Isto é, as propriedades acústicas do som são necessárias, mas não suficientes para explicar a capacidade, uma vez que as alturas não estão organizadas em função da distância linear entre as frequências. O fato de que as escalas musicais estão organizadas por meio de alturas não distantes entre si, e de que são essas as relações que determinam o padrão característico de tensão e relaxamento da música, assim criando expectativas em relação à sequência musical, é, para o autor, um argumento a favor da organização cognitiva – e não somente acústica – do espaço tonal. (LERDAHL, 2001). A experiência com determinado idioma musical é responsável pela criação de representações mentais e pelo processamento dessas relações, de maneira semelhante à qual o gerativismo explica a aquisição da linguagem.

## **Hipótese do Recurso de Integração Sintática Compartilhado**

No que diz respeito à natureza das relações cognitivas e neurais entre linguagem e música, o neurocientista cognitivo Patel (2003, 2008), baseado em estudos de neuroimagem, procura encontrar, no substrato neural, uma capacidade já tida como abstrata: a sintaxe.

Estudos em neuropsicologia mostram casos de amusia congênita ou adquirida<sup>8</sup> em que o indivíduo não apresenta nenhum tipo de déficit no que se refere à sintaxe linguística. No caso contrário, de afasia sem amusia, apenas aqueles que possuíam formação musical e/ou eram músicos profissionais não tinham suas capacidades sintáticas musicais afetadas (para uma revisão desses estudos, ver PERETZ; COLTHEART, 2003; PATEL, 2003). Estudos em neuroimagem, ao contrário, mostram evidências de sobreposição cerebral entre as capacidades sintáticas na música e na linguagem (PATEL *et al.*, 1998; FEDORENKO *et al.*, 2009; GIBSON *et al.*, 1998). Como, então, resolver esse aparente paradoxo?

Para lidar com esse impasse, Patel (2003, 2008) elabora a Hipótese do Recurso de Integração Sintática Compartilhado (SSIRH, *Shared Syntactic Integration Resource Hypothesis*, em inglês), segundo a qual as áreas do cérebro que provêm recursos para

---

<sup>8</sup> Assim como a afasia refere-se a uma deficiência na compreensão e produção linguística, a amusia refere-se a uma deficiência em que o indivíduo não consegue compreender música, por exemplo, perceber notas fora do tom. Pode ser causada por lesão cerebral (amusia adquirida), ou pode ser produto de uma condição genética (amusia congênita).



integração sintática são as mesmas áreas de processamento de linguagem e música. As evidências de sobreposição de processamento linguístico e musical sintático são explicadas por meio da sobreposição de áreas cerebrais onde a integração sintática ocorre. De acordo com essa hipótese, as redes neurais responsáveis por prover os recursos para a integração sintática e, portanto, para o processamento sintático, estão sobrepostas, enquanto as redes neurais responsáveis pela representação mental das estruturas linguísticas e musicais estão localizadas em partes diferentes do cérebro.

Por processamento sintático, Patel (2008) refere-se à maneira como o cérebro organiza os sons – que são elementos perceptualmente discretos – em estruturas hierárquicas. O autor justifica a sua comparação partindo do pressuposto de que tanto a música quanto a linguagem envolvem a produção e percepção de sons discretos, que são organizados de maneira complexa. Entretanto, uma divisão é feita entre o que é processamento sintático e o que é representação mental. As representações se referem ao conhecimento adquirido e armazenado no cérebro – que seria realizado em regiões diferentes, segundo a SSIRH. Ou seja, as operações cognitivas para o processamento das capacidades são as mesmas, enquanto as representações são de naturezas diferentes.

Essa abordagem destaca-se por vir como uma alternativa não só empírica, como também de cunho neurocognitivo, ao estudo da relação entre música e linguagem, associando argumentos de teorias cognitivas e discussões acerca das representações mentais da música e da linguagem a argumentos mais concretos, provenientes de estudos de neuroimagem da neurociência e da neuropsicologia. Entretanto, tanto a música quanto a linguagem continuam a ser tratadas como entidades psicológicas e abstratas.

## Hipótese da modularidade musical

A hipótese da modularidade musical, que adere ao conceito de modularidade formulado por Jerry Fodor (1983), em seu livro *The Modularity of Mind*, tem como premissa que “a música é uma faculdade cognitivamente única e distinta em termos evolutivos<sup>9</sup>” (PERETZ; COLTHEART, 2003, p. 688, tradução minha). Segundo essa hipótese, um módulo exclusivo à música representa um sistema mental de processamento de informação exclusiva ao *input* musical.

Também baseados em estudos de neuroimagem, como Patel (2003, 2008), e a partir da hipótese de Fodor de que a mente possui módulos específicos para funções e capacidades cognitivas específicas, Peretz e Coltheart (2003) apresentam uma arquitetura funcional modular da música, responsável pelo seu processamento por meio de diferentes submódulos e fluxos de informação. Da mesma maneira como a linguagem possui, na visão modular, subsistemas específicos à fonologia e ao léxico, o módulo musical

---

9 No original, “music is a cognitively unique and evolutionary distinct faculty”.

teria, nesta abordagem, submódulos específicos para o processamento de aspectos característicos à música.

Essa perspectiva é apoiada por estudos neuropsicológicos de indivíduos com afasia ou amusia, resultados de lesão cerebral. De um lado, observam-se casos em que a capacidade musical é afetada, enquanto a linguística se mantém intacta; de outro, verifica-se o contrário, casos em que a capacidade linguística é prejudicada, enquanto a musical permanece ilesa. Para Peretz e Coltheart (2003), essas são as evidências mais persuasivas de que o processamento de música e de fala não acontecem no mesmo sistema.

Entretanto, vale lembrar que a especialização neural não é requisito para um módulo musical. Segundo o modelo proposto pelos autores, sua propriedade fundamental é a especificidade do domínio. Diz-se que um mecanismo possui a propriedade de especificidade de domínio quando ele lida com processos altamente especializados, seja exclusivamente a um determinado *input* ou de maneira mais eficiente que outros mecanismos. Casos de lesão cerebral que provocam deficiências musicais sem afetar as habilidades linguísticas do indivíduo, bem como deficiências musicais que afetam o processamento do *pitch*, mas não o processamento do ritmo na música, e vice-versa, apontam que o componente responsável pela codificação tonal tem especialização neural (PERETZ; COLTHEART, 2003). No entanto, como aponta Peretz (2004), ainda não é possível apontar regiões cerebrais específicas ao processamento musical, mas defendem que, até o momento, estudos mostram que a música é processada em várias e vastas regiões neurais, localizadas em ambos os hemisférios, com uma especialização no hemisfério direito no processamento do *pitch* (PERETZ; HYDE, 2003; ZATORRE *et al.*, 2002 apud PERETZ, 2004).

A partir disso, Peretz (2009) sugere que a especificidade de um domínio possa ser resultado de mecanismos mais gerais da cognição ou de recursos compartilhados (como propõe a SSIRH), que se especializaram e se modularizaram com a experiência. A interação entre a música e a linguagem, por sua vez, pode ser explicada pela competição por mecanismos mais gerais da cognição, como atenção e memória.

## **Memória Sonora de Curto Prazo (ASTM)**

A proposta de Bigand *et al.* (2014) se contrapõe às ideias e hipóteses propostas na GTTM, na SSIRH e na hipótese da modularidade musical, estudadas até o momento; enquanto estas explicam as capacidades linguística e musical por meio de mecanismos de ordem mais abstrata, aquela se baseia em experimentos e explicações comportamentais e neurofisiológicas para explicá-las, ao mesmo tempo em que questiona a adequação de se postular uma “sintaxe” musical tal qual a sintaxe linguística.

Nesse estudo, uma série de experimentos a respeito dos mecanismos sintáticos envolvidos no processamento musical foram realizados, mas usando um modelo diferente para explicar os fenômenos, a saber, o modelo da Memória Sonora de Curto Prazo (ASTM, *Auditory Short-Term Memory*, em inglês). Esse modelo funciona como uma espécie de “ouvido artificial”, que calcula os níveis de estabilidade das alturas dentro de um contexto tonal. O que o modelo se propõe a fazer é explicar a percepção sonora por meio somente das propriedades acústicas do som e do seu armazenamento na memória sonora. Essas simulações têm como objetivo verificar quanto do processamento sintático musical pode ser explicado por um modelo ASTM, sem recorrer a influências computacionais abstratas ou representações mentais de natureza sintática.

Os resultados dos testes sugerem que o processamento de estruturas de natureza sintática na música pode não depender de representações mentais ou mecanismos abstratos, como aqueles envolvidos na linguagem. A música, ao contrário, seria processada somente pelas propriedades acústicas do *input* e pelas informações acústicas armazenadas na memória sonora, simulada pelo modelo ASTM. Bigand *et al.* (2014) apontam que essa conclusão gera implicações também na concepção de sintaxe musical, em relação à sintaxe linguística, levando-os a reavaliar até onde o processamento musical pode ser comparado ao processamento linguístico.

### ***A joint speech, vista a partir do paradigma dinâmico da cognição***

Do outro lado do espectro teórico que recortamos neste artigo está a perspectiva dinâmica da cognição, afastando-se das abordagens anteriores ao entender que a cognição humana é mais complexa do que o modo como as vertentes cognitivas e neurocognitivistas a concebem. O dinamicismo (para algumas referências, ver PORT; van GELDER, 1995; van GELDER, 1998; KELSO, 1995; THELEN; SMITH, 1994) vem como uma alternativa ao cognitivismo, procurando resolver algumas “questões mal resolvidas” na Ciência Cognitiva a respeito da noção de representação (HASELAGER, 2004).

Tal paradigma concebe o mundo e a mente que nele está situada de forma mais complexa que os modelos tradicionais da cognição, em que os fenômenos investigados são inseridos em um “mundo cristalino”, cujos limites estão claramente definidos, tal como em um jogo de xadrez (VARELA; THOMPSON; ROSCH, 1991). Não é o caso em uma abordagem dinâmica – como a ação de dirigir em uma cidade (para seguir a metáfora dos autores) –, em que os espaços não são pré-concebidos nem claramente delimitados; pelo contrário, esse mundo não termina em nenhum momento, e para conhecê-lo, é necessário ancorar-se em um conhecimento comum e no acúmulo de experiências.

Imaginemos agora uma partida de futebol: todas as arquibancadas do estádio lotadas, a bola em jogo e os torcedores gritando em uníssono, sejam palavras de apoio aos jogadores, sejam ofensas aos oponentes. Essa é uma cena familiar a muitos; se nunca

presenciada pessoalmente antes, ao menos já vista na televisão. Imaginemos também outra cena comumente vista – nas ruas ou nos jornais: uma multidão espalhada pelas ruas de uma cidade, de todas as idades, carregando bandeiras, cartazes e, na voz, um único grito de protesto. Podemos imaginar também uma terceira cena, talvez ainda mais familiar – menos agressiva, mas não menos veemente: fiéis em uma igreja ou em um templo, entoando juntos uma prece ou oração<sup>10</sup>.

O que há de comum entre essas três cenas? O mais notável é o fato de várias pessoas estarem falando a mesma coisa, ao mesmo tempo. Mas, afinal, o que há de curioso ou instigante nessas ações?

Esses são exemplos de um tipo de fala chamada por Cummins (2013, 2018) de *joint speech* (fala conjunta, em português). Esse é um termo “guarda-chuva” que se refere a uma fala conjunta, em que não há transmissão de uma mensagem de um falante a outro, mas sim a entoação de uma mesma mensagem por vários falantes, ao mesmo tempo. Como define o próprio autor: “A fala conjunta é a fala produzida por duas ou mais pessoas que enunciam a mesma coisa ao mesmo tempo<sup>11</sup>” (CUMMINS, 2018, p. 16, tradução minha). Essa definição empírica permite delimitar e categorizar atividades que, do ponto de vista antropológico e etnográfico, se mostram mais difíceis de serem definidas.

Como os exemplos das torcidas esportivas, protestos e orações ilustram, é um tipo de fala usado para fins coletivos, isto é, para manifestar uma identidade coletiva – seja uma adesão a um time, a uma ideologia ou reivindicação, ou a uma religião. Esse fenômeno é extremamente produtivo em diversas culturas humanas, e bem comum, mas pouco estudado.

Além da sua ubiquidade, outras características se fazem notáveis: (i) não há distinção entre falante e ouvinte<sup>12</sup>; (ii) a repetição é um traço canônico desse tipo de fala, de tal modo que a ausência de repetição também pode caracterizar um tipo de expressão da fala conjunta; (iii) a performatividade também a define, uma vez que precisa ser enunciada (*uttered*)<sup>13</sup> para que ela exista; (iv) por fim, a fala conjunta compartilha

---

10 Sugiro o *site* do autor para exemplos audiovisuais do fenômeno, em <http://jointspeech.ucd.ie/index.php/examples/>.

11 No original, “Joint speech is speech produced by two or more people who utter the same thing at the same time”.

12 Quando se fala que não há distinção entre falante e ouvinte, Cummins (2018) refere-se ao fato de que, na fala conjunta, os interlocutores não se alternam nos papéis de falante e ouvinte, como descrevem as teorias da conversação (GALEMBECK, 2010).

13 A respeito da performatividade, Cummins (2018) faz referência à noção de enunciação presente na Teoria dos Atos de Fala (AUSTIN, 1962).

muitas características com a música, e os limites entre linguagem e música se tornam praticamente invisíveis (CUMMINS, 2018).

Em termos acústicos, alguns aspectos são de particular interesse: por causa da repetição, as irregularidades rítmicas da fala se tornam mais ordenadas, uma vez que as batidas associadas com o início de sílabas fortes se tornam mais regulares ao longo do tempo; e o contorno da frequência fundamental (F0) se torna mais melódico, pois a prosódia dos versos repetidos se torna mais estilizada. Ou seja, os limites entre a música e a fala começam a se tornar menos claros quando se trata desse fenômeno.

Dessa maneira, a fala conjunta insere-se na discussão a respeito da relação entre a música e a linguagem justamente porque pode ser vista como o elo entre as duas capacidades. Cummins (2013) defende que a relação entre elas não é acidental; pelo contrário, propõe que sejam analisadas como polos de um mesmo contínuo, organizado pelo grau de coordenação das ações. Esse contínuo hipotético se inicia, à esquerda, pela música, passando pela canção, cântico (*chant*), fala conjunta, conversação, monólogo e fala silenciosa. É no centro, com o cântico e a fala conjunta, que os limites entre a música e a fala se misturam.

Para que possamos compreender não só esse contínuo, como também a música e a linguagem por uma perspectiva dinâmica, é preciso levar em conta outra definição de “comunicação”, diferente da tradicionalmente proposta pela Linguística. Cummins (2013), assim, propõe dois modelos de comunicação. Na concepção tradicional – o modelo do tubo –, a comunicação é a passagem de mensagens codificadas, e dá-se como um processo de etapas justapostas. No modelo da dança, para o qual a natureza coordenada da comunicação é importante, a comunicação se dá por meio de um padrão de comportamento coordenado, resultado da interação entre os sons da fala e os movimentos, comportamentos e experiências dos falantes. A grande diferença entre esses dois modelos trata da divisão feita entre o que é externo e o que é interno. Enquanto o modelo do tubo assume uma diferença entre o que é abstrato, interior, e um domínio físico, exterior, o modelo da dança não faz tais distinções. Outra diferença refere-se ao fato de que o primeiro modelo foca somente sobre o individual, ao contrário do segundo, que abrange também modos de fala que só podem ser compreendidos em função da atividade de outros.

Nesse contínuo, a coordenação da ação aumenta de um polo a outro. A propriedade coordenativa da ação se dá quando a ação de um participante é influenciada pelas ações do outro, e vice-versa, fazendo com que os participantes engajem em um processo contínuo de causalidade. Percebe-se, portanto, que a ação é mais coordenada quanto mais próxima do polo musical, por ser a música uma atividade mais coordenada que a fala. O ponto de encontro entre a música e a fala é justamente a fala conjunta e os cânticos (*chants*), em que as características de ambas as atividades se misturam, e os limites entre elas se tornam mais difusos.

## Uma expressão da fala conjunta: a fala sincronizada

A fala sincronizada, por sua vez, é um tipo de fala conjunta obtida em experimento. Comparando a fala à performance musical, a fala sincronizada se assemelha à ação de tocar em conjunto, pois a primeira demanda ao falante que restrinja sua fala de maneira semelhante a um músico que toca em conjunto com outros. A música, no entanto, providencia um tempo ideal a partir do qual se pode medir a variação. Resta saber se na fala há fatores temporais inerentes a ela, semelhantes ao tempo ideal da performance musical, ou se se trata de fatores aplicados opcionalmente pelo falante conforme a situação comunicativa.

Experimentos (CUMMINS, 2002, 2003) procuraram verificar se a fala obtida em sincronia com outra pessoa pode ser diferenciada de outras formas de fala, como a leitura individual sem restrições, ou a fala obtida em sincronia com uma gravação. Os resultados demonstraram que a sincronia entre a fala de dois indivíduos é fácil e rapidamente obtida sem nenhum tipo de treinamento prévio, apesar de os falantes não apresentarem uma prosódia estilizada e previsível, como em orações e recitações. Os falantes concordam praticamente por completo em relação às pausas quando estão lendo em pares, mas adicionam mais pausas quando estão lendo individualmente. A sincronização é de tal modo obtida que a diferença de medida entre dois pontos correspondentes de dois sinais de fala (*lag*) não dura mais do que 40 ms, em média, aumentando para 60 ms após pausas, no início das sentenças.

Isso suscita a questão de como essa sincronia é adquirida, afinal, para que a sincronia se faça possível, os falantes precisam ajustar-se uns aos outros. Para isso, precisam tornar suas falas menos idiossincráticas e mais previsíveis. Duas hipóteses são então levantadas: ou os falantes voltam a valores intrínsecos de segmentação e duração das sílabas, ou, de maneira dinâmica, os sujeitos compartilham sinais ou pistas que permitem que a parte seguinte seja previsível. Para resolver a questão, Cummins (2002) comparou a performance dos sujeitos na leitura em sincronia com dois tipos de gravação, o primeiro com as gravações de leitura individual e a segunda com um dos canais da fala sincronizada em par. Os resultados mostraram *lags* menores na condição de fala sincronizada (média de 30 ms), em relação à condição de sincronização com uma gravação (média de 56 ms), sugerindo que o que está em questão na fala sincronizada é um processo de acomodação ao co-falante, e não um processo de imitação do outro. A concordância em relação à colocação das pausas também mostra que os falantes procuram produzir uma fala menos expressiva e variável.

Se os falantes, portanto, são capazes de entrar rapidamente em sincronia e formar um nível de organização diádica, sugere-se que estão agindo – metaforicamente – como osciladores sincronizados dentro de um sistema dinâmico, na medida em que as atividades dos falantes são mutuamente influenciadas. De fato, os falantes parecem

agir como tal, pois sua relação não é de líder/seguidor, isto é, não há no sistema um controlador central, mas sim uma coincidência temporal no comportamento sequencial dos falantes.

Sendo assim, a visão dinâmica da cognição mostra-se pertinente para explicar tais fenômenos, pois a ausência de periodicidade nessa fala levanta questões a respeito do fator que possibilita o processo. Segundo essa perspectiva, que integra a produção da fala à sua percepção, a ação de falar envolve não só os movimentos realizados pelo falante e o som produzido por ele, mas também o som percebido. Dessa maneira, a fala sincronizada (e a fala conjunta, como um todo) é resultado da integração da fala de um indivíduo à fala do outro, em um processo de influência contínua, em que os falantes tornam-se, assim, conjuntamente responsáveis pela fala resultante e pelo ato comunicativo. Trata-se, afinal, do modelo da dança, em oposição ao modelo do tubo (CUMMINS, 2013).

Avaliando a influência de sinais acusticamente modificados sobre o processo de sincronização, os resultados de Cummins (2009) sugerem que não são informações específicas que permitem o fenômeno, mas sim que são informações contínuas ao longo do sinal que contribuem para a sincronização entre falantes. Esse resultado, em particular, favorece um modelo que vê a sincronização como integração (*entrainment*), em que as informações acústicas possuem um papel análogo às informações visuais (e às vezes, táteis) usadas na movimentação corporal coordenada entre indivíduos. Isto é, a troca contínua de informações acústica entre os falantes permite a coordenação dos movimentos que estão envolvidos na produção da fala.

Relacionado a essa noção de *entrainment*, Cummins (2009) também introduz à discussão a noção de *affordance*, ao afirmar que o ritmo constitui um *affordance* para o movimento. Conceito ligado ao trabalho de Gibson (1986) e à psicologia ecológica e traduzido por Raposo de Medeiros (2009) como viabilização, diz respeito à propriedade do ambiente que é relevante ao movimento potencial de um organismo, ou seja, é o produto das relações entre as estruturas físicas do ambiente e o intelecto (percepção e abstração) dos seres. Essa visão estabelece uma relação entre o sinal acústico e o potencial para movimento de um ouvinte incorporado (*embodied*).

Ainda tratando-se do ritmo em atividades sincronizadas, seu papel nesse processo é questionado quando pensamos que não há periodicidade na fala, em oposição a atividades periódicas como o canto e a música, em geral. Sobre isso, destaca-se aqui o estudo de Raposo de Medeiros e Cummins (2014), cujo questionamento diz respeito à influência do tipo de texto e suas particularidades rítmicas sobre a tarefa de sincronização, ou seja, se a periodicidade de um texto pode facilitá-la. A sincronização entre falantes foi testada por meio da fala e do canto de diferentes tipos de texto, que apresentam estruturas rítmicas variadas. Foram selecionados um texto em prosa (ausência de uma estrutura rítmica regular), uma parlenda (presença de uma estrutura rítmica regular), uma canção

do gênero *rock* (ritmo simples) e uma canção do gênero *samba* (ritmo complexo)<sup>14</sup>. Em resumo, os resultados mostraram que a sincronização ocorreu de maneira semelhante em todos os casos, corroborando os estudos anteriores, de que a sincronização da fala não requer periodicidade.

## Discussão final

Este trabalho teve como objetivo realizar um percurso sobre diversas abordagens das Ciências da Cognição, desde aquelas que veem a capacidade musical e linguística como abstrações de um comportamento, regras e categorias armazenadas na mente, até o dinamicismo, que passa a vê-las como sistemas emergentes de processos diversos, cognitivos e sociais.

Acertadamente, Haselager (2004) sugere que o uso excessivo do conceito de representação no cognitivismo pode levar ao seu esvaziamento. Se antes a Ciência Cognitiva procurou uma alternativa às explicações demasiadamente materialistas do Behaviorismo, agora faz-se necessário buscar alternativas a um modelo excessivamente mentalista. As abordagens do paradigma dinâmico da cognição se mostram como uma boa alternativa, uma vez que deixam de lado o representacionismo em função da importância da interação corporal com o ambiente.

Pensar a capacidade linguística como um sistema adaptativo complexo é concebê-la como uma capacidade emergente do funcionamento cognitivo geral e da interação social, e não apenas como um módulo mental, como ela é tradicionalmente vista. Nessa concepção, a linguagem forma um todo complexo cujas partes se retroalimentam, constantemente mudando com o tempo e a experiência (ALBANO, 2012).

A natureza da relação entre a música e a linguagem, seus paralelos e suas diferenças lançam um questionamento: se não é necessário uma abordagem que enfatize a importância da interação do corpo com o ambiente para explicar comportamentos inteligentes como esses. A abordagem dinâmica, então, se destaca, ao conceber a cognição como corporificação e movimento, levando em conta o princípio de que o todo não corresponde à soma das partes, mas sim à integração coordenada e dinâmica delas. Em oposição à ideia de categorização, traz as noções de continuidade e gradiência. Diferentemente da visão cognitivista, o dinamicismo incorpora ao seu modelo a experiência e o contexto em que o indivíduo está inserido, tratando da relação da experiência e do comportamento dos sons – sejam eles linguísticos ou musicais – e gestos dos falantes de maneira integrada.

---

14 A respeito do ritmo do *rock* e do *samba*, e de suas classificações, ver Raposo de Medeiros (2014).



Tendo isso em vista, a fala conjunta e a fala sincronizada se destacam tanto no estudo comparativo das capacidades quanto dentro da abordagem dinâmica porque são fenômenos em que a interação dos falantes com o ambiente é central. Esses fenômenos questionam justamente a noção de representação mental e sua necessidade para explicar a cognição. Abordar a música e a fala como polos de um mesmo contínuo, como Cummins (2013) faz, e não como capacidades opostas ou distantes, permite explicar tanto o compartilhamento de propriedades e características quanto diferenças entre elas.

O percurso teórico realizado neste trabalho teve também como propósito apresentar o campo ainda pouco explorado que é o estudo da fala conjunta e da fala sincronizada. Seu estudo interessa não só à investigação da natureza da relação entre a música e a linguagem e as características compartilhadas por elas, mais do que isso, ele abre espaço para discutir os conflitos entre o individual e o coletivo, e como esses são resolvidos nesse tipo de ação. Ademais, diversos diálogos são possibilitados pela área; além dos estudos musicais, com a psicologia, o estudo do movimento e, até mesmo, o estudo da enunciação e dos atos de fala.

## REFERÊNCIAS

- ALBANO, E. C. *Da fala à linguagem tocando de ouvido*. São Paulo: Martins Fontes, 1990.
- ALBANO, E. C. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. *Revista da Abralin*, n. 2, p. 1-30, 2012.
- AUSTIN, J. L. *How to do things with words*. Cambridge, Massachusetts: Harvard University Press, 1962.
- BIGAND, E.; POULIN-CHARRONAT, B. Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, n. 100, p. 100-130, 2006.
- BIGAND, E. *et al.* Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, v. 8, art. 94, jun. 2014.
- CHOMSKY, N. *Syntactic Structures*. London: Mouton, 1957.
- CHOMSKY, N. *Aspects of the Theory of Syntax*. Cambridge: M.I.T. Press, 1965.
- CHOMSKY, N. *Language and Mind*. New York: Harcourt, Brace & World, 1968.

CHOMSKY, N. *The Logical Structure of Linguistic Theory*, New York: Plenum Press, 1975.

CUMMINS, F. On synchronous speech. *Acoustic Research Online*, n. 3, v. 1, p. 7-11, 2002. Disponível em: <https://asa.scitation.org/doi/10.1121/1.1416672>. Acesso em: 20 set. 2018.

CUMMINS, F. Practice and performance in speech produced synchronously. *Journal of Phonetics*, v. 31, p. 139-148, 2003.

CUMMINS, F. Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics*, n. 37, p. 16-28, 2009.

CUMMINS, F. Joint speech: The missing link between speech and music? *Percepta*, n. 1, v. 1, p. 17-32, 2013.

CUMMINS, F. *The Ground From Which We Speak: Joint speech and the collective subject*. Cambridge: Cambridge Scholars Publishing, 2018.

FEDORENKO, E. *et al.* Structural evidence in language and music: Evidence for a shared system. *Memory & Cognition*, v. 37, n. 1, p. 1-9, 2009.

FODOR, J. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

GALEMBECK, P. T. O turno conversacional. In: PRETI, D. (org.). *Análise de textos orais*. v. 1. 7. ed. São Paulo: Humanitas, 2010.

GIBSON, J. J. *The Ecological Approach to Visual Perception*. Psychology Press, 1986.

HASELAGER, W. F. G. O mal-estar do representacionismo: sete dores de cabeça da Ciência Cognitiva. In: FERREIRA, A.; GONZALEZ, M. E. Q.; COELHO, J. G. (ed.). *Encontros com as Ciências Cognitivas*, v. 4. São Paulo: Coleção Estudos Cognitivos, 2004. p. 105-120.

KELSO, J. *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, MA: MIT Press, 1995.

KOLINSKY, R. *et al.* Processing interactions between phonology and melody: Vowels sing but consonants speak. *Cognition*, v. 112, p. 1-20, 2009.

JACKENDOFF, R. Parallels and Nonparallels between Language and Music. *Music Perception*, v. 26, n. 3, p. 195-204, 2009.

JACKENDOFF, R.; LERDAHL, F. The capacity for music: What is it and what's special about it? *Cognition*, v. 100, p. 33-72, 2006.

LERDAHL, F. Tonal Pitch Space. *Music Perception*, v. 5, p. 315-350, 1988.

LERDAHL, F. *Tonal pitch space*. New York: Oxford University Press, 2001.

LERDAHL, F.; JACKENDOFF, R. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press, 1983.

MCDERMOTT, J.; HAUSER, M. The origins of music: innateness, uniqueness and evolution. *Music Perception*, v. 23, n. 1, p. 29-59, 2005.

MILLER, G. A. *Language and Communication*. New-York-London: McGraw-Hill, 1951.

NETTL, B. *The Study of Ethnomusicology: Thirty-one Issues and Concepts*. Urbana and Chicago: University of Illinois Press, 1983.

NEWELL, A.; SIMON, H. A. *Human Problem Solving*. Oxford, England: Prentice-Hall, 1972.

PATEL, A. Language, music, syntax and the brain. *Nature Neuroscience*, v. 6, n. 7, p. 674-681, jul./2003. Disponível em: <https://www.nature.com/articles/nn1082>. Acesso em: 20 set. 2018.

PATEL, A. *Music, language and the brain*. New York: Oxford University Press, 2008.

PATEL, A. *et al.* Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, v. 10, n. 6, p. 717-733, 1998.

PERETZ, I. Brain Specialization for Music: New Evidence from Congenital Amusia. In: PERETZ, I.; ZATORRE, R. (ed.). *The Cognitive Neuroscience of Music*. New York: Oxford University Press, 2004.

PERETZ, I. The nature of music from a biological perspective. *Cognition*, n. 100, p. 1-32, 2006.

PERETZ, I. Music, language and modularity framed in action. *Psychologica Belgica*, v. 49, n. 2/3, p. 157-175, 2009.

PERETZ, I.; COLTHEART, M. Modularity of music processing. *Nature Neuroscience*, v. 6, n. 7, p. 688-691, jul. 2003.

PORT, R.; van GELDER, T. *Mind as motion*. Cambridge, MA: MIT Press, 1995.

RAPOSO de MEDEIROS, B. *Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro*. 2002. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

RAPOSO de MEDEIROS, B. Ritmo na língua e na música: o elo possível. *Música em Perspectiva*, v. 2, n. 2, p. 45-63, 2009.

RAPOSO de MEDEIROS, B. A organização temporal na fala e no canto: estudo com foco na sincronização. *Percepta*, v. 1, n. 2, p. 27-50, 2014.

RAPOSO de MEDEIROS, R.; CUMMINS, F. Speech and song synchronization: A comparative study. *Proceedings of the 7th International Conference on Speech Prosody*, Dublin, Ireland, p. 748-751, 2014.

RAPOSO DE MEDEIROS. B.; CABRAL, J. P. Acoustic distinctions between speech and singing: Is singing acoustically more stable than speech? *Proceedings of the 9th International Conference on Speech Prosody*, Poznan, Poland, p. 542-546, 2018.

SUNDBERG, J. *The science of the singing voice*. Northern Illinois University Press, Dekalb, 1987.

THELEN, E.; SMITH, L. B. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press, 1994.

van GELDER, T. The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, n. 21, v. 5, p. 615-628, 1998.

VARELA, F. J.; THOMPSON, E. T.; ROSCH, E. *The Embodied Mind: Cognitive Sciences and Human Experience*. Cambridge, MA: MIT Press, 1991.