

- Shin, H. & Kim, J. (2011). Music Therapy on Anxiety, Stress and Maternal-Fetal Attachment in Pregnant Women During Transvaginal Ultrasound. *Asian Nursing Research*, 5(1), 19-27.
- Tseng, Y., Chen, C. & Lee, C. (2010). Effects of listening to music on postpartum stress and anxiety levels. *Journal of Clinical Nursing*, 19, 1049-1055.
- Ventura, T. Gomes, M. S. & Carreira, T. (2012). Cortisol and anxiety response to a relaxing intervention on pregnant women awaiting amniocentesis. *Psychoneuroendocrinology*, 37, 148-156.
- Winnicott, D. W. (1956). A preocupação Maternal Primária. In *Da Pediatria à Psicanálise* (pp. 399-405). Lisboa: Imago Editora.

VOZ CANTADA E VOZ FALADA NO GERMINAR DA COMUNICAÇÃO HUMANA: ESTUDO SOBRE VOCALIZAÇÕES DE BEBÉS NO SEGUNDO ANO DE VIDA

João Reigado
Helena Rodrigues

Nota introdutória: variações sobre fios de voz

A música cantada e a linguagem falada estão presentes em todas as sociedades. Com maior ou menor definição dos seus limites, a manifestação destas formas de comunicação humana tem acompanhado a evolução humana ao longo dos tempos. Diferentes perspectivas teóricas alimentam o interesse acerca das semelhanças e diferenças entre estas formas de comunicação, sendo fascinante o debate acerca da sua génese evolutiva (para uma revisão sobre o tema ver Wallin, Merker, & Brown, 2000). Uma das perspectivas mais divulgadas reivindica o termo *musilanguage* – proposto por Brown (2000) –, defendendo que música e linguagem terão evoluído como especializações de um ancestral comum.

Se se verificar, no âmbito do desenvolvimento musical, que a ontogénese repete a filogénese, talvez possamos admitir que a primeira produção vocal da criança compartilha características musicais e linguísticas. Ou seja, talvez possamos ver refletido no desenvolvimento vocal da criança, um modelo *musicolinguístico* próximo do proposto por Brown relativo ao desenvolvimento vocal na evolução humana.

Dar corpo a esta hipótese obriga-nos a explorar as semelhanças e diferenças entre o cantar e o falar em idade precoce, encontrando elementos caracterizadores de uma e de outra forma de comunicação. Ou seja, a estabelecer uma análise no terreno da produção vocal numa etapa do desenvolvimento infantil caracterizada pela constante transformação, pela intencionalidade comunicativa e pela aquisição de competências de distinta natureza. Talvez canto e fala se alimentem do mesmo respirar; talvez canto e fala se mobilizem a partir de idênticos atos motores; talvez canto e fala brotem da mesma necessidade de contacto com os outros. A génese será a mesma, mas talvez canto e fala se estruturarem e, eventualmente, se diferenciem, em função dos diálogos estabelecidos com esses mesmos outros e com o seu espaço cultural.

Apresente-se, então, o primeiro conjunto de interrogações: serão as vocalizações do bebé apenas precursoras dos exercícios linguísticos iniciais ou poder-se-á considerar uma fase elementar de vocalização com partilha de características musicais e linguísticas? Quando um adulto afirma que aquele bebé começou por cantar antes de falar está a dizer o quê? O que é que os seus ouvidos escutaram? Qualquer adulto interpretaria da mesma forma aquilo que escutou? O ato de “cantar” na primeira infância, observado por muitos Pais, que parecem não ter dúvidas em reconhecer o significado de “cantar” e “falar” (pelo menos, quando se trata do seu próprio bebé) será apenas produto da imaginação babada dos pais ou será mais do que isso?

Um conjunto de estudos (a que à frente se fará referência) parece reforçar a ideia da existência precoce de um tipo de comportamento vocal em resposta a motivos musicais, por parte do bebé. Nesses estudos fica por esclarecer o que é que do comportamento vocal infantil é biológico e independente da contextualização e o que é que é fruto da modelização do adulto e da cultura em que a criança está inserida. Mas teremos algum dia respostas satisfatórias para este tipo de questões que, de resto, radicam nas mais antigas questões filosóficas da humanidade?

Alimente-se, pois, esse debate, percorrendo os estádios de desenvolvimento geral, musical e linguístico da criança, inquietando-nos com duas questões interdependentes: em que período do desenvolvimento infantil é possível encontrar traços distintivos entre produção vocal falada e produção vocal cantada? Que características, partilhadas e não partilhadas, constituem a base das vocalizações consideradas musicais e as consideradas linguísticas?

Perceber a partir de quando estamos na presença de comportamentos vocais, de algum modo relacionados com a música cantada que as crianças escutam, incentiva-nos a sistematizar um conjunto de práticas educativas que possam ir ao encontro das capacidades musicais que as crianças já dominam. Mais do que isso, colocará à disposição do professor e do educador elementos fundamentais para a elaboração de materiais musicais mais válidos e diretamente orientados para as crianças de que vão cuidar.

Estimulação, predisposição e precursores do canto no desenvolvimento vocal de bebés

Vários autores reportam diferenças na interação dos bebés com objetos e com pessoas (Brazelton & Cramer, 1989; Murray, 1991; Trevarthen, 1990), mostrando que na interação com pessoas há, desde o início, uma intenção comunicativa (Trevarthen, 1990; Gomes-Pedro, 1985), na qual são visíveis gestos e movimentos pré-linguísticos. Esta interação, que se estabelece desde o início da

vida, é carregada de afetos, tendo-lhe sido apontada, por vários investigadores, traços de musicalidade.

De entre o conjunto de aquisições de âmbito comunicativo observadas nos primeiros meses de vida do bebé, destaca-se o uso da voz. A produção de sons pelo bebé atrai o interesse da mãe, reforçando a comunicação entre ambos (H. Papousek, 1996; M. Papousek, 1996; Van Puyvelde *et al.*, 2010).

Por outro lado, a prosódia da fala materna parece desempenhar um estímulo fundamental para as primeiras vocalizações dos bebés, a par de padrões de comportamento partilhados que, de acordo com Papoušek (M. Papoušek, 1996), são comuns entre culturas. Note-se, porém, que os aspectos musicais universais presentes nas interações pré-verbais entre pais e bebés não escondem o cariz genuíno e exclusivo que também está presente na comunicação entre aqueles. Isto é, para além da universalidade dos elementos musicais presentes naquelas interações, cada díade parece encontrar, nas palavras de Rodrigues (p. 7, 2005), “um código próprio de comunicação”.

A altura dos sons – qualidade que nos permite percecioná-los como *agudos* ou *graves* – parece estar também envolvida na capacidade de os bebés reproduzirem balbucios com entoação, antes da produção de fala significativa. É relevante que, para H. Papoušek (1995), não seja “o conteúdo lexical mas a entoação métrica que assume a importância nas nossas primeiras conversações com os bebés” (p. 172). Ou seja, mais importante do que a palavra em si, a entoação da produção vocal com que nos dirigimos ao bebé (entoação em sentido lato, englobando aspectos relativos à altura, ao ritmo, ao timbre, à expressividade) proporciona-lhe uma das primeiras experiências sociais. Como defende Rodrigues (2004), “[...] observando a interação que sucede naturalmente entre o bebé e quem dele cuida, somos levados a constatar todo um conjunto de comportamentos marcado por formas básicas de comunicação caracterizadas por grande expressividade e musicalidade” (Rodrigues, 2004, p. 9).

A representação mental de melodias nos primeiros anos de vida da criança parece também enfatizar esta atração pela entoação da linha melódica. A atração pelo contorno melódico, isto é, pelo padrão de sequências de subidas e descidas da altura dos sons, como observa Trehub (Trehub, 2003; Trehub *et al.*, 1997), manifesta-se na primeira etapa de desenvolvimento da perceção sonora e musical dos bebés e é anterior à perceção de intervalos melódicos e notas individuais. Tal significa que os bebés processam, em primeiro lugar, aspectos holísticos dos padrões sonoros e só mais tarde atentam a aspectos mais detalhados. É possível também que os mecanismos de agrupamento auditivo¹

¹ De acordo com a Teoria da Gestalt o cérebro agrupa os elementos percetivos (auditivos, visuais ou de outra natureza) de acordo com as características que possuem entre si. Estes agrupamentos são formados a partir da aplicação de vários princípios (leis) que regem a perceção humana. Assim,

observados em adultos (Bregman, 1994) funcionem desde muito cedo no desenvolvimento humano.

A primeira forma de comunicação que o bebé descobre como despertando o interesse dos pais será o choro e, pouco depois, o sorriso. Apesar da comunicação não se estabelecer exclusivamente através destas formas de expressão, estas serão a alavanca dos primeiros balbucios. Para Costa (2003), esta é uma etapa fascinante que prepara a produção de linguagem explícita. Reconhecer se um determinado balbucio corresponde a uma verdadeira palavra é possível se se perceber a intenção comunicativa da criança. Ou seja, para Costa, “quando o mesmo som passa a ser utilizado num contexto específico, com uma intenção comunicativa determinada e exclusivamente na presença daquilo que a palavra significa, então pode ser interpretado como uma palavra” (Costa, p. 85).

A estrutura destas primeiras palavras corresponde a sílabas formadas geralmente por apenas uma vogal (ou por uma consoante e uma vogal – CV), como no caso de “na” para a palavra “não”. Rapidamente ganham forma as primeiras sequências bem formadas de unidades CV – a criança começa a produzir cadeias de sílabas do tipo “mamama” ou “papapa”. Tipicamente entre os 9 e os 18 meses pode observar-se a produção das primeiras palavras por entre o balbucio. Esta fase pode conter, simultaneamente, unidades identificáveis como elementos significativos, isto é, pertencentes ao vocabulário de uma determinada língua, e sílabas que não são reconhecíveis em termos lexicais.

Apesar de as primeiras vocalizações dos bebés terem sido amplamente investigadas em contextos linguísticos, poucas tentativas foram avançadas relativamente à discussão do fenómeno relacionando-o com as origens do desenvolvimento musical.

Para alguns autores, as vocalizações de conteúdo musical estão presentes nos primeiros anos de vida da criança (Moog, 1976; Fox, 1983; Gordon, 1990; Welch, 1994, 2006; Dowling, 1999). Estes autores fundamentam esta convicção a partir da observação de bebés e crianças em contextos ecológicos, argumentando que as suas vocalizações contêm características musicais semelhantes às que são típicas da cultura musical desse mesmo ambiente.

As primeiras produções vocais parecem basear-se, sobretudo, em linhas melódicas descendentes, isto é, terminando num som mais grave do que o som inicial. Parece também consensual que a vocalização cantada consista, inicialmente, na utilização de sons isolados, maioritariamente vocálicos (isto é, sons sem consoantes).

o agrupamento auditivo é a organização de elementos sonoros entre si de modo a formarem todos coerentes. Na Psicologia da Música, a teoria da Gestalt teve grande impacto no estudo do ritmo, ao nível da formação de padrões temporais.

Tafuri e Villa analisaram as vocalizações espontâneas de bebês, num estudo longitudinal que seguiu famílias e os seus bebês durante os primeiros 6 anos de vida (Tafuri & Villa, 2002). O projeto *inCanto* envolveu 119 mães, inicialmente no 6.º ou 7.º mês de gravidez, e os seus filhos. As futuras mães participaram em sessões semanais de atividades musicais até ao momento do nascimento do bebé e em sessões de orientação musical para recém-nascidos a partir do final do primeiro mês após o nascimento. Como componente do programa de educação musical do projeto, as mães foram encorajadas a realizar um conjunto de práticas diárias em casa durante a gravidez e, mais tarde, durante a comunicação estabelecida com os seus bebês. Tais práticas incluíam canções simples, jogos de movimento rítmico e audição de música gravada. Foi-lhes ainda pedido que elaborassem um diário acerca da natureza e duração das atividades musicais realizadas em casa e que anotassem e gravassem periodicamente as reações vocais dos seus bebês. Foram também gravadas as vocalizações de um conjunto de bebês pertencentes a um grupo de controlo, sem exposição a atividades musicais.

Uma das vertentes do projeto *inCanto* considerou a descrição dos elementos musicais presentes nas vocalizações produzidas pelos bebês do estudo, entre os 2 e os 8 meses. Observou-se que os bebês vocalizaram alguns glissandos (ascendentes e descendentes) e intervalos de 3.^a, 4.^a e 5.^a. Tafuri e Villa registaram também a ocorrência de cadências e progressões melódicas, isto é, sequências de sons que, no primeiro caso, transmitem um sentido de resolução de uma ideia/frase musical e que, no segundo, dizem respeito a estruturas maiores, eventualmente mais próximas das melodias cantadas por adultos (Tafuri & Villa, 2002). As gravações obtidas revelaram ainda maior número de combinação de sons e produção de vocalizações de maior duração do grupo de bebês face ao grupo de controlo (Tafuri & Villa, 2002).

Numa outra vertente do projeto, procurou-se encontrar uma relação entre o envolvimento musical precoce dos bebês e a sua posterior produção vocal de canto espontâneo. A análise dos diários e das gravações áudio das vocalizações indicou que entre os 2 e os 3 anos as crianças já exercitam a criação livre de canções inventadas, cuja análise levou à consideração de três categorias: frases, monólogos e canções (Tafuri, 2003). Nestas produções inventadas a autora observou uma presença clara de intervalos e padrões melódico-rítmicos próprios da cultura musical ocidental. Tafuri conclui que, desde que inseridas em contextos que estimulem as atividades musicais, as crianças nesta idade possuem já os mecanismos fisiológicos, perceptivos e cognitivos para a invenção de canções.

A musicalidade na raiz da comunicação: elementos acústicos na comparação entre produção vocal em contexto de voz cantada e voz falada

Embora alguns dos autores acima mencionados tenham arriscado comparações entre as modalidades da voz falada e cantada, nenhum é suficientemente claro relativamente às semelhanças e diferenças acústicas presentes nas produções vocais das crianças.

Pelo contrário, em estudos acerca da voz adulta, são conhecidas as características acústicas que diferenciam voz cantada e voz falada. Neste âmbito, a manipulação das vogais foi identificada como uma característica importante para a caracterização das duas formas de expressão vocal (Sundberg, 1987, 1999; Scotto Di Carlo, 2005; Patel, 2008). Nomeadamente, verificou-se que na passagem do registo falado para o cantado, a duração das vogais aumenta, enquanto a parte consonântica de uma sílaba tende a diminuir (Scotto Di Carlo & Autesserre, 1992).

Poder-se-á perguntar se esta característica acústica, importante na diferenciação entre voz cantada e voz falada em idade adulta, ocorre também nas vocalizações produzidas na infância.

Tratar-se-á de atender a um aspecto fundamental da qualidade da vibração vocal – o *timbre* – muitas vezes excluído do estudo das produções vocais infantis, uma vez que não se refere a uma dimensão sonora unitária. A sua análise resulta da relação de todas as propriedades acústicas que não influenciam diretamente a altura e a intensidade do som (Henrique, 2002).

O timbre parece ser percecionado de forma multidimensional, sendo as dimensões temporal e espectral fundamentais no estudo comparativo entre voz cantada e voz falada. A dimensão temporal refere-se à evolução da amplitude de um som ao longo do tempo, enquanto a dimensão espectral diz respeito à distribuição das diferentes frequências que o compõem e às amplitudes relativas dessas mesmas frequências.

Ao contrário da importância da altura dos sons, não parece que existam sistemas musicais organizados exclusivamente a partir de contrastes tímbricos. É na linguagem que a dimensão tímbrica assume toda a sua importância, uma vez que a fala consiste, fundamentalmente, num sistema organizado de contrastes tímbricos.

Dos sons da fala, a vogal será o mais musical, contendo uma frequência fundamental (F0) clara (correspondente à altura do som que é escutado), e uma estrutura harmónica rica, isto é, um grande número de harmónicos que normalmente são múltiplos exatos de F0 (por exemplo, F0 = 120Hz, F1 = 240 Hz, F2 = 360 Hz, etc.) A estrutura acústica de uma vogal pode variar bastante face ao contexto de produção. Tal como se referiu anteriormente, Scotto Di Carlo (Scotto Di Carlo, 2005; Scotto Di Carlo e Autesserre, 1992) mostra que na passagem da

voz falada para a voz cantada, a duração média das vogais aumenta, enquanto que a das consoantes diminui ligeiramente. Assim é porque as vogais constituem um suporte ideal para o fluxo melódico, enquanto que as consoantes agem como barreira à continuidade melódica. Este aspecto reforça a importância do estudo do timbre na comparação entre as produções vocais faladas e cantadas.

Uma outra dimensão sonora importante na comparação entre voz cantada e voz falada é a *altura* do som. Esta é uma característica psicológica cujo correspondente físico é a frequência da onda sonora, medida em Hertz (Henrique, 2002)². Uma vez que todas as culturas possuem algum tipo de manifestação musical baseada na diferenciação de alturas de som (Patel, 2008), esta dimensão (e não outras como, por exemplo, a intensidade do som) parece ser fundamental para a criação de um qualquer sistema musical organizado.

A altura desempenha também um papel muito importante no contexto de línguas tonais (isto é, línguas em que a mesma palavra pode adotar significados distintos, dependendo da altura dos seus sons: por exemplo, o mandarim). Em línguas não-tonais a altura continua a ser uma característica importante (de outro modo teríamos uma fala monocórdica sem lugar para a interrogação ou a exclamação) mas as diferenças na altura dos sons numa mesma palavra não lhe conferem diferentes significados lexicais.

A este respeito, refira-se um estudo pioneiro de Mang (2001), através do qual se procurou verificar se falantes de línguas tonais e de línguas não-tonais exibiam um padrão de altura vocal distintivo. Mang observou o desenvolvimento vocal de um grupo de oito crianças inglesas e chinesas, durante 42 meses, tendo obtido a altura das suas vocalizações a partir de uma análise acústica. A autora comparou longitudinalmente as F0 vocais entre os 2 e os 6 anos de idade das crianças, verificando que as diferenças entre voz cantada e voz falada aumentaram com a idade nas crianças falantes de inglês, mas tal não sucedeu com as crianças cuja língua materna era o mandarim.

Ou seja, nas crianças de língua materna inglesa há uma maior distinção entre voz falada e voz cantada, enquanto que no mandarim o parâmetro altura vocal na produção cantada e falada não parece ser uma característica distintiva proeminente. Estes resultados vêm ao encontro do que seria de esperar, dado que o mandarim é uma língua tonal e, portanto, mais próxima do canto; não obstante, é interessante verificar que as vocalizações de crianças desta faixa etária exibem já as características das respetivas línguas maternas e da sua diferença relativamente à voz cantada.

² Henrique (2002) recomenda prudência quanto ao uso do termo altura para classificar um som, visto tratar-se de uma característica psicológica e não física. Tendo em conta a recomendação do autor, neste texto optou-se pela utilização desta expressão sempre que não estava implícita qualquer medição física do som e pela expressão frequência fundamental nos casos correspondentes à medição de uma grandeza física.

A maior parte das melodias musicais constrói-se em torno de um conjunto estável de intervalos melódicos, enquanto que as melodias linguísticas não. Tal permite, por exemplo, que as melodias musicais possam recorrer a um centro tonal, como anteriormente se aludiu. Patel (2005) considera que esta característica está na base daquilo que torna as melodias musicais tão eficazes. Pelo contrário, os contornos entoacionais de fala raramente são memorizados e cantarolados uma vez que a rede de relações entre os tons entoacionais de fala não é tão rica como a das melodias musicais, tornando-se esteticamente menos relevante.

De acordo com vários autores (Guimarães, 2007; Henrique, 2002; Patel, 2005; Cook, 2001), um dos parâmetros mais importantes na caracterização melódica da voz falada é a determinação da F0, a partir da frequência de vibração das cordas vocais. Uma característica evidente da F0 da fala é o seu movimento, para cima e para baixo, através de glissandos muito rápidos (Xu & Sun, 2002), sem que nenhuma diferenciação ordenada de tons ou semitons possa ser percebida. Esta linha melódica da fala, bastante ondulante, contrasta com a sequência de tons de altura definida das melodias de instrumentos musicais e, certamente, evidenciará igualmente diferenças relativas à voz cantada.

No que concerne à caracterização melódica da voz cantada, um aspecto que parece ser comum tanto à música de tradição ocidental, como à proveniente de outras culturas, é a predominância de intervalos pequenos entre os sucessivos sons (Huron, 2006). Ou seja, parece observar-se uma tendência para que o movimento melódico se desenvolva por graus conjuntos.

Em resumo, estaremos perante a produção vocal falada ou cantada, em resultado da manipulação de parâmetros idênticos. Assim, o falar e o cantar representam categorias diferentes, mas a inter-relação profunda entre ambas, torna a sua separação difícil.

Desenvolvimento vocal na infância: análise acústica de vocalizações de bebés

O conjunto de estudos referidos amplia o conhecimento acerca do comportamento vocal musical das crianças em contextos ecológicos de observação. No entanto, os dados referidos não estabelecem uma relação entre o tipo de vocalizações produzidas pelos bebés e as canções cantadas/ouvidas pelas mães, sendo igualmente pouco claros os critérios usados para a classificação das produções vocais “inventadas” pelas crianças.

Por outro lado, poucas referências são feitas à comparação entre as produções vocais que ocorrem em contextos de natureza linguística e as que ocorrem em contextos de natureza musical que, no nosso entender, deverão ser consideradas.

Outro aspecto comum aos estudos atrás apresentados é a utilização de unidades de análise que poderão não ser adequadas ao estudo do desenvolvimento musical infantil. Esta é, aliás, uma característica comum em investigações sobre desenvolvimento musical na primeira infância, nas quais se descreve o comportamento musical infantil confrontando-o com o padrão adulto. Concretamente, numa análise musical de material produzido por adultos poderá fazer sentido utilizar critérios baseados na produção de intervalos (referindo, por exemplo, a produção de intervalos de 3.^a, 4.^a ou 5.^a); mas será que faz sentido utilizar exatamente os mesmos critérios quando se está a analisar produções vocais musicais que, do ponto de vista da cultura musical ocidental, são ainda incipientes? Isto é, do nosso ponto de vista, há que encontrar descritores específicos para caracterizar o desenvolvimento musical na infância tal como têm sido encontrados para o desenvolvimento linguístico, para descrever cada uma das suas etapas.

Assim, num estudo de Reigado (2007) em que se procurou comparar o desenvolvimento da voz cantada com o desenvolvimento da voz falada, adotou-se uma perspetiva de análise distinta das anteriormente descritas. Nesse sentido, procuraram-se aspectos acústicos que permitissem caracterizar as vocalizações de bebés, tendo-se encontrado dados que apontam para a existência de uma produção vocal específica para o contexto musical e, portanto, diferente do tipo de vocalização produzida em contexto linguístico.

Foram analisadas as vocalizações de 21 bebés, com idades entre os 9 e os 11 meses, produzidas no contexto de estimulação musical e linguística. Os bebés foram observados semanalmente, ao longo de quatro sessões de 20 minutos cada, num ambiente controlado mas próximo do que poderia ser um contexto ecológico de interação com bebés. Cada sessão compreendia a apresentação de duas situações: uma situação musical, em que o experimentador cantava, e uma situação linguística, em que o experimentador narrava poemas num estilo de voz próximo do chamado discurso dirigido às crianças³. As vocalizações dos bebés, produzidas durante os dois tipos de estimulação foram gravadas e posteriormente submetidas a análises auditivas (perceptivas) e acústicas.

Verificou-se que as vocalizações que os bebés produziam em contexto musical apresentavam características diferentes das produzidas em contexto lin-

³ De acordo com Fernald e Simon (1984), quando as mães falam com os seus filhos bebés utilizam um tipo de discurso com características próprias, designando-o por "infant-directed speech". Ao fazê-lo, recorrem a sons de altura mais aguda, oscilando vocalmente entre uma amplitude de alturas mais vasta do que na fala habitual com adultos. Outras características deste estilo de discurso dirigido às crianças prendem-se com a existência de pausas mais longas, enunciados mais curtos e maior número de repetições prosódicas. O "infant-directed speech" realça fortemente a prosódia da língua materna, atraindo mais facilmente a atenção dos bebés, quando comparado com a fala adulta convencional (Fernald & Kuhl, 1987).

guístico. Concretamente, as vocalizações produzidas face à estimulação musical registaram menor duração do que as obtidas perante a estimulação linguística. No que concerne à estrutura das vocalizações, constatou-se, face à estimulação musical, uma preponderância acentuada de vocalizações de sons isolados, enquanto que, relativamente às vocalizações produzidas em resposta à estimulação linguística, se observou um equilíbrio entre a ocorrência de sons isolados, de estruturas de duas ou mais sílabas e de contornos entoacionais. Foram também encontradas diferenças quanto à extensão vocal usada, observando-se que esta foi mais limitada no caso do contexto musical. Assim, nas respostas vocais ao estímulo musical, os bebés usam maioritariamente uma extensão limite de 9 meios-tons e em contexto linguístico uma extensão máxima de 25 meios-tons. Note-se, no entanto, que este facto poderá apenas revelar que os bebés adequaram a sua extensão vocal à extensão vocal dos estímulos apresentados.

Partindo do mesmo *corpus* de vocalizações, através de uma nova pesquisa verificou-se que a vocalização de sons isolados observada em Reigado (2007) dificultava o reconhecimento pelo adulto do contexto de vocalização (Rocha, 2007). Ou seja, a falta de um número mínimo de determinadas características sonoras impedia o adulto de constituir uma representação mental da vocalização dos bebés, que permitisse ligá-la inequivocamente a um contexto musical ou linguístico.

Este facto acabou por definir o problema de partida da investigação que adiante se reporta: a ocorrência de sons isolados estará condicionada pela idade das crianças (no estudo anterior os bebés participantes tinham entre 9 e 11 meses) ou será consequência dos estímulos utilizados?

Pretendeu-se então acompanhar a evolução da voz cantada durante o segundo ano de vida, confrontando características das vocalizações produzidas em duas condições distintas: conversa e canção.

Ou seja, era importante verificar se, no caso de existirem diferenças nas vocalizações das crianças produzidas em contexto linguístico ou musical, estas ocorriam não só porque a criança progredia em termos de desenvolvimento linguístico, mas também em termos da sua voz cantada.

Nesse sentido, foram estudadas as vocalizações dos bebés sob diferentes perspetivas: análise comparativa dos elementos segmentais presentes nas produções vocais obtidas nos contextos musical ou linguístico; análise comparativa das características melódicas e rítmicas, bem como o estudo dos índices de variabilidade de intervalos acústicos – *nPVI* (*normalized pairwise variability index*) para os núcleos, para os intervalos consonânticos e para intervalos silábico – e índice de variabilidade dos intervalos melódicos presentes numa vocalização – *MIV* (*melodic interval variability*); estudo longitudinal das vocalizações produzidas em contexto musical.

Um relato aprofundado deste trabalho pode ser encontrado em Reigado (2013).

Voz cantada e voz falada no segundo ano de vida: relato de um estudo

Na sequência do exposto foi conduzida uma investigação que pretendeu acompanhar o desenvolvimento vocal da criança durante o segundo ano de vida, estabelecendo comparações entre produção vocal em resposta à voz cantada e em resposta à voz falada. Apresentam-se de seguida os respetivos aspectos metodológicos.

Participantes

Participaram neste estudo doze crianças (7 raparigas e 5 rapazes), acompanhadas por um dos pais. No início do estudo todas as crianças tinham 12 meses de idade ($M = 11,5$). Os pais das crianças acederam em participar no estudo, durante todo um ano, ao longo do segundo ano de vida dos seus filhos, sem compensações monetárias. Como incentivo à participação dos pais e dos seus filhos, foi acordada a possibilidade de frequentarem um conjunto de sessões de orientação musical para pais e bebés, no final da fase de recolha de dados.

Materiais

Três canções (abaixo transcritas) foram compostas com o propósito de que as mesmas refletissem contrastes musicais, quer internamente, no que concerne à forma de organização das frases musicais, quer quando comparadas entre si, relativamente à tonalidade, à métrica e ao andamento.

Pato corredor
(Maior / binária) J.R.



$\text{♩} = 72$ Fine

9 D.C. al Fine

Barco a remos

(menorh / ternária)

J.R.



$\text{♩} = 140$

8

13

Dança do caracol

(dórico / binária)

J.R.



$\text{♩} = 120$

9

Figura 1. Canções sem palavras usadas no estudo

As sessões foram gravadas através de 3 microfones (DPA 4060 Hi-Sens), suspensos a partir do tecto da sala e do interface M-Audio ProFire™ 2626 (24bit/192kHz) com pré-amplificação (Octane™) e tecnologia JetPLL para a eliminação de jitter, conectados a um computador Apple, iMac 24" (2.8GHz/8GB SDRAM/640GB HD). Usou-se o programa Audacity 1.3.9-beta (<http://audacity.sourceforge.net/>) para segmentação e etiquetagem das amostras sonoras. No apoio à análise acústica e tratamento das amostras vocais foram utilizados os programas Praat 4.4.04 (Boersma & Weenink, 2006) e Prosogram (Mertens, 2004).

Procedimento

Foi efetuada a recolha de vocalizações ao longo do segundo ano de vida de 12 bebés participantes. Para tal, foram organizados encontros de 30 minutos entre o experimentador e o bebé, acompanhado por um adulto acompanhante (normalmente o Pai ou a Mãe), no espaço do Laboratório de Música e Comu-



Figura 2. Espaço da recolha de dados

nicação na Infância, do Centro de Estudos de Sociologia e Estética Musical, da Faculdade de Ciências Sociais e Humanas (ver fig. 2). Cada bebé foi sujeito a duas condições de estimulação: a condição Conversa e a condição Canção. Na condição Conversa o experimentador dirigia-se à criança falando; na condição Canção o experimentador dirigia-se à criança cantando, usando para tal uma sílaba sem significado linguístico (“bam-bam”). A condição Conversa baseava-se tanto em situações de interação linguística como em momentos de pausa, isto é, períodos em que o experimentador permanecia em silêncio e destinados à obtenção de vocalizações produzidas pelas crianças. A condição Canção contemplava três canções, interações musicais e pausas. As situações de interação musical destinavam-se à execução, pelo experimentador, de padrões musicais de cada canção, tentando-se uma analogia entre os padrões musicais executados e os padrões silábicos comumente usados em modelos de acesso lexical (ver, por exemplo, Rivera & Smith, 1997).

Os encontros referidos realizaram-se quinzenalmente entre os 12 e os 24 meses de vida dos bebés.

Foram gravadas as vocalizações obtidas durante o segundo ano de vida, selecionando-se para análise as ocorrentes aos 12, 15, 18, 21 e 24 meses.

Obteve-se um total de 662 produções vocais ($N = 662$), constatando-se que 449 ocorreram na condição Conversa e 213 ocorreram na condição Canção. A amostra final excluiu sons vegetativos (tais como sibilos, espirros, tosse e soluços), sons de esforço (tais como gemidos e grunhidos) e vocalizações obtidas em contexto de birra ou choro. Cada vocalização em análise correspondeu a todo o sinal acústico situado entre o final da vocalização estímulo, produzida pelo experimentador, e o final do último elemento vocalizado pela criança, imediatamente antes de novo estímulo vocal do experimentador.

Face às características do estudo acústico e à utilização do *software* Praat, versão 4.4.04 (Boersma & Weenink, 2006) foram ainda rejeitadas amostras cujo

sinal acústico era demasiado fraco e cuja duração era inferior a 1500 ms. Para a análise acústica das variáveis, recorreu-se a medições com base no espectrograma e forma de onda, cuja banda de detecção de F_0 foi estabelecida entre 40 a 800 Hz, com uma frequência de amostragem de 200 Hz.

Medidas

Com recurso ao programa Prosogram (Mertens, 2004) foram calculados os parâmetros prosódicos do sinal de voz, conseguindo-se uma segmentação em sílabas e pausas. Deste modo, foram depois determinados o início e o final das vogais usados para calcular a duração de cada vogal ou sequência de vogais e de consoantes, bem como a duração compreendida entre inícios de núcleos consecutivos.

Numa primeira fase, determinou-se a duração total de uma vocalização – *durvocal* – e o tempo de fonação (a duração de uma vocalização descontando pausas entre fonações) – *durfonal*. Contabilizou-se o número de núcleos por vocalização – *nnucleos* – e a duração média (em segundos) dos núcleos – *dnucleos*. Foi ainda apurada a velocidade da produção vocal – *ratenucleos* – calculada a partir da razão entre o número de núcleos e o tempo total de uma vocalização. Este primeiro conjunto de medidas correspondeu à caracterização da estrutura geral de cada vocalização, observando a sua duração, o número de elementos constituintes e a velocidade de articulação vocal nas duas condições. Ou seja, é o que se costuma designar como elementos segmentais de uma vocalização.

Um segundo conjunto de cálculos considerou a caracterização das vocalizações face à altura dos sons (através da determinação de F_0) e aspectos melódicos. Consideraram-se três parâmetros para esta variável, representando as frequências fundamentais mínima ($\text{min}F_0$), média ($\text{med}F_0$) e máxima ($\text{max}F_0$) de cada vocalização. A determinação destes três patamares de F_0 interessa na medida em que poderá revelar diferenças na extensão vocal da vocalização em resposta a uma e outra condições de estimulação.

Em termos melódicos, no estudo efetuou-se uma análise do contorno melódico, próxima da percepção da variação da altura dos sons e baseada no sinal acústico. A medida *intradinF0* representa a percentagem do tempo em que a variação da frequência fundamental dentro de um núcleo excedeu o limiar de glissando, isto é, em que a variação da frequência foi entendida como movimento ascendente ou descendente. A variável *interabsF0* representa a soma das variações absolutas de F_0 entre núcleos sucessivos.

Descrição da amostra

O Quadro 1 apresenta a distribuição do número de vocalizações recolhidas por condição experimental e sexo da criança.

	CONVERSA		CANÇÃO		Total
	n	%	n	%	
Rapaz	116	55,24	94	44,76	210
Rapariga	333	73,67	119	26,33	452
N	449		213		

Quadro 1. Número Total de Vocalizações: Distribuição por Sexo e Condição Experimental

O número total de vocalizações é, para a condição Conversa, superior ao dobro do registado na condição Canção. O número total de vocalizações é bastante mais elevado no caso de bebés raparigas (N = 452) do que em bebés rapazes (N = 210). Tal parece ser o resultado de uma maior assimetria de comportamento de produção vocal dos bebés raparigas, relativamente às duas condições (Conversa = 333; Canção = 119). O mesmo facto não se verifica no comportamento vocal dos bebés rapazes cuja distribuição pelas duas condições é mais equilibrada (Conversa = 116; Canção = 94). Não obstante, não deixa de ser interessante verificar que, para qualquer dos sexos, a condição Conversa proporciona maior número de produções vocais.

A percentagem de vocalizações tem, por isso, um comportamento desigual quando analisada tendo em conta o sexo e a condição experimental. O teste do qui-quadrado permite rejeitar a independência entre o sexo e a condição experimental [$\chi^2(1) = 21,491$, ($p < 0,001$)], admitindo-se a existência de associação entre a condição experimental em que ocorre a vocalização e o sexo da criança.

Apresentação de resultados

Significância estatística das diferenças entre as vocalizações em função da condição experimental

Podemos observar no Quadro 2 que a condição experimental – *Conversa vs. Canção* – influencia significativamente a duração da vocalização (*durvocal*) (F(1,660) = 9.77, $p < .05$), as variáveis *maxF0* (F(1,516) = 4,36, $p = 0,037$) e *medF0* (F(1,516) = 4,80, $p = 0,029$) e muito significativamente o tempo de fonação (*durfonal*) (F(1,660) = 4.23, $p < .01$), o número de núcleos (*nnucleos*) (F(1,660) = 12.20, $p < .01$) e a duração dos núcleos (*dnucleos*) (F(1,660) = 8.77, $p < .01$).

	CONVERSA		CANÇÃO		F
	M	DP	M	DP	
durvocal	1,52	1,46	1,26	0,99	5.54*
durfonal	0,98	0,77	0,81	0,58	8.19**
nnucleos	5,48	4,47	4,30	3,11	12.2**
dnucleos	0,09	0,04	0,10	0,06	8.77**
ratenucleos	6,13	4,89	6,89	8,58	2.09
minF0	258,33	151,56	243,77	147,63	0,99
maxF0	563,49	151,71	531,99	164,37	4,36*
medF0	405,73	114,81	381,26	115,72	4,8*
intradinF0	9,84	13,70	10,70	15,60	0,41
interabsF0	33,80	48,30	30,30	36,30	0,65

* $p < .05$ ** $p < .01$

Quadro 2. Valores Médios, Desvios-Padrão e Análise de Variância das Variáveis Acústicas de Acordo com a Condição Experimental

Pretendeu-se analisar a influência quantitativa de cada variável independente – sexo, condição experimental e idade – sobre as variáveis em estudo nas vocalizações das crianças. Conduziu-se uma análise de regressão linear, cujos coeficientes apurados para o efeito das variáveis independentes acima indicadas, vêm expressos no Quadro 3.

Pela análise do Quadro 3, constata-se que o facto de o bebé ser rapariga ou rapaz tem pouco efeito sobre as características em estudo nas vocalizações. Por outro lado, a condição experimental parece exercer um efeito significativo sobre as variáveis *durvocal*, *durfonal*, *maxF0* e *medF0* ($p < 0,05$) e muito significativo sobre as variáveis *nnucleos* e *dnucleos* ($p < 0,01$). Relativamente à tendência de variação das variáveis dependentes face ao efeito da condição experimental, observa-se que quando a condição é Canção os valores de *durvocal*, *durfonal*, *nnucleos*, *maxF0* e *medF0* diminuem, enquanto que *dnucleos* aumenta.

A análise de regressão linear mostra ainda que a idade da criança tem também um efeito significativo sobre as variáveis *durfonal*, *dnucleos*, *maxF0* e *interabsF0* ($p < 0,05$) e muito significativo sobre a variável *ratenucleos* ($p < 0,01$). Quanto à tendência de variação destas variáveis observa-se que à medida que a idade da criança aumenta, os valores de *durfonal* e *dnucleos* aumentam, enquanto que diminuem os valores de *ratenucleos*, *maxF0* e *interabsF0*.

	SEXO		Condição Experimental		Idade	
	(Rapariga=0, Rapaz=1)		(Conversa=0, Canção=1)			
	B	p	B	p	B	p
durvocal	-0,139	0,219	-0,244	0,033*	-0,008	0,574
durfonal	-0,017	0,774	-0,143	0,021*	0,019	0,010*
nnucleos	-0,152	0,663	-1,099	0,002**	0,045	0,275
dnucleos	0,002	0,573	0,012	0,002**	0,001	0,038*
ratenucleos	-0,432	0,419	0,600	0,267	-0,179	0,005**
minF0	-6,56	0,657	-11,894	0,430	1,361	0,431
maxF0	-13,994	0,357	-33,203	0,032*	-4,483	0,012*
medF0	-8,46	0,455	-23,548	0,041*	-0,757	0,566
intradinF0	-0,451	0,749	1,054	0,462	0,074	0,654
interabsF0	-0,471	0,915	-4,477	0,321	-1,024	0,048*

*p < .05 **p < .01

Quadro 3. Coeficientes da Regressão Linear do Efeito do Sexo, Condição Experimental e Idade da Criança sobre as Variáveis Acústicas

Efeito da idade sobre o desenvolvimento da voz cantada

A regressão linear mostrou que, para além da condição experimental, também a idade da criança exercia efeito sobre algumas das variáveis em análise nas suas vocalizações. Analisa-se, agora, este efeito em detalhe, observando-se longitudinalmente os dados recolhidos em 5 momentos do 2.º ano de vida da criança.

As próximas figuras exibem a tendência de variação dos elementos segmentais (ver fig. 3), da altura dos sons e dos aspectos melódicos (ver fig. 4) analisados nas vocalizações produzidas na condição Canção ao longo do segundo ano de vida.



Figura 3. Variação dos elementos segmentais das vocalizações produzidas na condição Canção. Duração da vocalização (durvocal), tempo de fonação (durfonal) e número de núcleos (nnucleos) (esquerda); duração média de um núcleo (dnucleos) (centro); velocidade de produção vocal (ratenucleos) (direita) aos 12, 15, 18, 21 e 24 meses de vida.

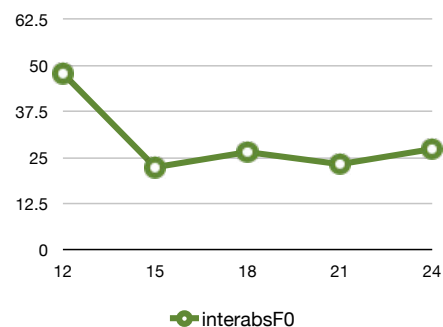


Figura 4. Variação da altura dos sons e dos aspectos melódicos analisados nas vocalizações produzidas na condição *Canção*. Valores de minF0, medF0 e maxF0 (esquerda); porcentagem de tempo em que a variação da frequência foi entendida como movimento ascendente ou descendente (centro); variação absoluta de F0 entre núcleos - *interabsF0* (direita) aos 12, 15, 18, 21 e 24 meses de vida.

Tal como o esperado, a duração da vocalização (*durvocal*) e o tempo de fonação (*durfonal*) apresentam uma tendência de variação semelhante. Pode-se observar que, apesar do número de núcleos (*nnucleos*) aumentar no período entre os 18 e os 21 meses, a duração da vocalização (*durvocal*) e o tempo de fonação (*durfonal*) diminuem. Tal deve-se ao facto de também a duração dos núcleos (*dnucleos*) diminuir neste período. Existe também um aumento paralelo das variáveis *durvocal*, *durfonal* e *dnucleos* entre os 21 e os 24 meses. A velocidade de produção vocal (*ratenucleos*) diminui entre os 12 e os 18 meses. É a partir deste ponto que a evolução desta variável depende diretamente da tendência de variação do número de núcleos (*nnucleos*), aumentando entre os 18 e os 21 meses (aumenta também *nnucleos*) e diminuindo entre os 21 e os 24 meses (diminui também *nnucleos*). De facto, o teste *post-hoc* HSD de Tukey para comparações múltiplas confirmou diferenças significativas para o número de núcleos (*nnucleos*) entre os 15 e os 18 meses ($p = .001$) e entre os 15 e os 24 meses ($p = .023$).

A porcentagem do tempo em que a variação da frequência é entendida como movimento (*intradinF0*) é significativamente diferente ao longo do tempo (ver fig. 2, ao centro). Concretamente, para a variação ocorrida entre os meses 21 e 24, o teste *post-hoc* HSD de Tukey confirmou também diferenças significativas ($p = .007$).

Relativamente à variável *interabsF0*, o período 12-15 meses ($p = 0.039$) regista médias que são significativamente diferentes. Ou seja, a variação absoluta de F0 entre núcleos de uma vocalização regista valores médios distintos em cada um destes dois períodos.

Discussão de resultados

A análise conduzida clarifica a existência de diferenças nas vocalizações de bebés produzidas nas duas condições em estudo (*Conversa* e *Canção*). Do ponto de vista da análise comparativa dos elementos segmentais, os resultados estão

em linha com os encontrados no estudo precedente (Reigado, 2009), confirmando a ocorrência de vocalizações mais longas na condição *Conversa*.

Nesse estudo anterior, a observação de produções vocais mais longas na sequência do estímulo linguístico revelou-se como um elemento distintivo entre as vocalizações produzidas numa e noutra situação. Propôs-se, na altura, que este facto fosse o resultado de uma maior exposição diária à estimulação linguística por parte dos bebés (que será, por assim dizer, o seu meio privilegiado de expressão). Ou seja, a menor familiarização com estímulos musicais implicaria uma assimetria de desenvolvimento da vocalização musical, em comparação com a evolução do balbucio linguístico.

No entanto, ficará por investigar quais as características que influenciavam diretamente a duração de uma vocalização, nas duas condições de estímulo. Com a realização deste segundo estudo, foi então possível constatar uma significativa superioridade do número de núcleos na condição *Conversa*, determinando a ocorrência de vocalizações mais longas neste contexto. Este facto pode ser explicado por uma maior familiarização das crianças com as características da linguagem verbal. De facto, vários estudos mostram que as crianças produzem, por volta do 2.º ano de idade, articulações claras e uma grande quantidade de séries de sílabas variadas (Boysson-Bardies, 2001; M. Papoušek, 1996). Mas pode também ver-se neste comportamento da criança, a reprodução das características do estímulo que o tornam diferente do estímulo musical. Ou seja, uma vez que o estímulo falado (condição *Conversa*) apresenta mais sílabas do que o estímulo cantado (condição *Canção*) (tanto mais que, recorde-se, a canção foi vocalizada numa sílaba apenas) é natural que a criança vocalize com maior número de sílabas naquele contexto.

Não deixa, no entanto, de ser interessante verificar que a duração média dos núcleos é superior na condição *Canção*. Ou seja, apesar das produções vocais ocorridas na condição *Canção* serem mais curtas, a duração dos núcleos é superior, denunciando aquilo que parece ser um prolongamento sistemático das vogais, quando as crianças vocalizam naquele contexto. Estes resultados confirmam as observações de Dowling (1999), que verifica que, por volta dos 2 anos, há uma maior sustentação de vogais nas vocalizações em resposta à condição *Canção*, do que relativamente à condição *Conversa*.

Estão também em linha com as observações de Scotto Di Carlo (2005) sobre a inteligibilidade da voz cantada adulta, quando constata que a sílaba sofre um prolongamento substancial no canto, o que leva a um aumento na duração relativa das vogais e diminuição nas consoantes. Para aquela autora, tal sucede porque as vogais são um suporte melódico ideal enquanto as consoantes afetam a qualidade do *legatto*, interrompendo a linha melódica. É surpreendente que, neste estudo, os resultados apurados apontem para um comportamento diferenciador semelhante na produção vocal infantil.

Do ponto de vista da análise das características melódicas das vocalizações, encontraram-se também algumas diferenças. Constata-se, em primeiro lugar, que os valores de F_0 encontrados para qualquer dos patamares de medição – $minF_0$, $maxF_0$ e $medF_0$ – foram superiores na condição Conversa. Estes resultados confirmam as observações de Reigado, Rocha e Rodrigues (2011), verificando-se, em ambos os estudos, que as crianças restringem a extensão vocal em contexto musical de forma a aproximá-la da extensão do estímulo cantado pelo adulto.

Refiram-se agora os resultados apurados acerca do comportamento da linha melódica, tanto nos núcleos, como na vocalização completa.

Uma tendência observada que merece reflexão surge da análise da variação de F_0 ao longo de toda a vocalização. Repare-se que, quando se observa a variação tonal considerando a sequência de intervalos melódicos da vocalização ($interabsF_0$), apuram-se valores superiores na condição Conversa aos encontrados na condição Canção. Este facto dá conta de uma vocalização que, na condição Canção, obedece a uma linha melódica com poucos saltos, ou seja, que recorre a tons próximos em termos de altura.

Este tipo de comportamento foi já observado por Gordon (2000), estando aliás na base dos conselhos que preconiza na forma de atuar com a criança que progride do estágio de resposta aleatória para o de resposta intencional da audição preparatória:

“No estágio 3 da audição preparatória apenas se devem cantar para as crianças os padrões tonais em tonalidades maior e menor harmónica que se movimentam diatonicamente (por graus de uma escala), porque estes são os mais característicos das vocalizações e dos sons da fala da própria criança.” (p. 71)

Estas observações parecem também convergir com as de Patel (2008), quando verifica a ocorrência, nas melodias musicais, de um maior número de intervalos pequenos.

A verificação deste tipo de comportamento em idade infantil é significativa. Não é de espantar que os intervalos de menor amplitude sejam mais fáceis de executar, devido à imaturidade e constrangimentos do aparelho vocal das crianças.

Assim, os dados relatados parecem ir ao encontro do que se pode observar em sessões de música para bebés e de iniciação musical com crianças pequenas. Nestas, a ocorrência de produções vocais que contemplam intervalos diatónicos é comum, sobretudo quando os bebés e as crianças são confrontados com melodias que terminam repentinamente, num ponto em que soam inconclusivas. Frequentemente, as crianças tendem a completar estas melodias can-

tando sons que são próximos, em termos de altura, do último som cantado e deixado em suspenso. Neste estudo, as crianças tenderam também a integrar o fluxo sonoro, produzindo sons vocais que completavam a coesão musical dos pequenos fragmentos musicais apresentados.

Este é um aspecto que permite uma ligação óbvia às estratégias usadas na percepção adulta do fluxo sonoro, tal como descrito por Bregman (1994). De acordo com o autor, a compreensão de mensagens auditivas obedece a um conjunto de regras que, uma vez violadas, dificultam a percepção das mesmas. Por exemplo, a proximidade de elementos sonoros vizinhos, como aqui se discute, parece ser um importante factor para a compreensão do fluxo musical (*op. cit.*, p. 401). Para se fazer compreender, linguística e musicalmente, o adulto tenderá também a utilizar estas regras na produção vocal que dirige a terceiros.

Em consonância com o que Bregman observa relativamente à percepção adulta, o presente estudo mostra que a criança parece também procurar estratégias que visam a compreensão da sua produção vocal em contexto musical, pelos sujeitos que lhe estão próximos.

Às conclusões de Bregman – relativas ao facto da proximidade de elementos sonoros vizinhos ser um importante factor para a compreensão do fluxo musical na situação perceptiva do adulto – há que acrescentar o facto de o bebé tender a vocalizar sons de alturas próximas entre si. Perante estes dados podemos argumentar que o bebé está dotado de um equipamento de produção vocal que permite que a sua mensagem seja percebida pelos adultos que lhe estão próximos. Podemos também especular que as características de percepção e de produção vocal andam a par e passo desde a primeira infância no que diz respeito à proximidade de elementos sonoros vizinhos.

Ao observar-se longitudinalmente os elementos segmentais e os aspectos melódicos das vocalizações produzidas face à condição Canção, não foram encontradas tendências inequívocas de crescimento (positivo ou negativo) das variáveis que caracterizam aqueles parâmetros. O ato vocal infantil é um fenómeno psicológico complexo, que requer a aquisição e coordenação de capacidades a vários níveis (perceptivo, físico, articulatorio e cognitivo), muito diferente de processos evolutivos lineares (como o crescimento físico, por exemplo). Efetivamente, o desenvolvimento da voz cantada parece assentar em reorganizações sucessivas de comportamentos.

De facto, a observação longitudinal demonstrou que a idade da criança influencia tanto a duração da vocalização como o número e duração dos núcleos. Assiste-se também a um efeito significativo da idade da criança sobre o tempo de oscilação de F_0 num núcleo e sobre a variação de F_0 entre núcleos consecutivos de uma vocalização. Razões de ordem física e articulatoria estarão na base destas observações. Como se sabe, o trato vocal das crianças passa por transformações de ordem morfológica durante os primeiros anos (Kuhl &

Meltzoff, 1996). Assim, a vogal, que será, em oposição à consoante, a parte tonal da sílaba, aumenta ao longo do tempo, tendo como consequência o aumento progressivo da duração das vocalizações cantadas. Este resultado evidencia a manipulação da vogal como um aspecto fundamental no desenvolvimento das vocalizações cantadas.

Em suma, esta investigação procurou examinar aspectos da aquisição da voz cantada no segundo ano de vida, considerando elementos segmentais e aspectos melódicos.

A análise das vocalizações revelou diferenças nas condições Conversa e Canção relativamente às seguintes características: duração das produções vocais, duração média da fonação, número médio de núcleos por vocalização, duração média de cada núcleo, valores de frequência fundamental – F0 e variações da frequência fundamental (F0) entre núcleos sucessivos de uma vocalização. A observação longitudinal confirmou ainda o efeito significativo da idade da criança sobre a duração da vocalização e o número e duração dos núcleos presentes numa vocalização.

Algumas questões ficam ainda por responder e obrigam a uma análise mais detalhada em futuros estudos. A primeira diz respeito ao aprofundamento das relações tonais entre estímulo e resposta, procurando verificar se existem graus tonais que são produzidos mais frequentemente. A segunda centra-se na importância que os registos filmados poderão ter para a análise, na medida em que permitem cruzar os processos de aquisição e desenvolvimento da voz cantada, com aspectos do movimento corporal, da expressão facial e, por exemplo, do direcionamento do olhar. Uma terceira questão diz respeito ao modo como se poderão conciliar instrumentos de análise, encontrando um fio condutor entre a análise técnica e a apreciação humana, especializada e dita não especializada.

Tanto quanto é possível saber, a partir da revisão da literatura, nenhum estudo abordou a produção vocal infantil utilizando como parâmetros de análise os elementos segmentais e melódicos usados no presente estudo. Assim, esta metodologia pode ser promissora para o acompanhamento do desenvolvimento musical das crianças, permitindo-nos ampliar o nosso conhecimento acerca da produção vocal em idades anteriores e posteriores à considerada nesta investigação. Esta metodologia poderá ser igualmente útil na análise de vocalizações produzidas por diferentes participantes (criança *versus* adulto; bebé *versus* mãe; etc.) e envolvendo diferentes estímulos musicais (vocal *versus* instrumental).

Outra potencialidade da investigação sobre a aquisição da voz cantada poderá ser a de conciliar instrumentos de análise, plataformas de observação. Em concreto, propõe-se que o estudo das primeiras vocalizações cantadas encontre um fio condutor entre a análise técnica e a apreciação humana,

especializada e dita não especializada. Retomando uma das questões iniciais, importa compreender o que está a suceder quando Pais e músicos sustentam que o (seu) bebé cantou. Sabemos que a construção do ato comunicativo assenta na partilha de significados; por isso, as observações e interpretações dos Pais (ou outros que estão próximos do bebé no dia a dia) não podem ser ignoradas.

Quando um bebé vocaliza “pa” poucas dúvidas existem que terá dito a palavra “pai”, ao invés de “pau”, “pão” ou “paralelepípedo”. Os Pais recriam o balbucio, idealizando-o como uma palavra que ainda não o é. Corporizam-na no pensamento da criança, que mais tarde a verbalizará. No fundo, esta idealização das aquisições linguística e musical que os Pais demonstram, de forma tão espontânea, é transversal a todos os comportamentos infantis em desenvolvimento. Porque a criança só interage com quem lhe devolve compreensão, com quem a contextualiza, lhe dá referências e a motiva para arriscar ser compreendida.

Precisamos de entrar na teia das vocalizações dos bebés, nesse espaço babélico, povoado de cores, sons e contornos e que, antes de fazer parte do mundo da música ou do da linguagem, expressam a razão pela qual se expressam – a sua veemente necessidade de comunicar. Nasce-se para con(tactar).

Referências

- Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.4.04) [Computer software]. Retrieved February 2006, from <http://www.praat.org/>
- Brazelton, T. B., & Cramer, B. (1989). *A relação mais precoce*. Lisboa: Terramar.
- Bregman, A. (1994). *Auditory Scene Analysis*. Massachusetts: MIT Press.
- Brown, S. (2000). The “musilanguage” model of music evolution. In N. L. Wallin, B. Merker, and S. Brown (Eds.) *The Origins of Music* (pp. 271-300). Cambridge, MA: MIT Press.
- Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The Psychology of Music* (pp.215-264). San Diego: Academic Press.
- Cook, P. (2001). Pitch, Periodicity, and Noise in the Voice. In P. Cook (Ed.), *Music, Cognition, and Computerized Sound. An Introduction to Psychoacoustics*. Cambridge, Massachusetts: MIT Press.
- Costa, J. & Santos, A. (2003). *A falar como os bebés – o desenvolvimento linguístico das crianças*. Lisboa: Editorial Caminho.
- Dowling, W. (1999). The development of music perception and cognition. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 603-625). San Diego: Academic Press.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers’ speech to newborns. *Developmental Psychology*, 20(1), 104-113.

- Fox, D. (1983). The pitch range and contour of infant vocalizations. (Doctoral dissertation, The Ohio State University, 1982). *Dissertation Abstracts International*, 43, 2588A.
- Gluschkof, C. (2002). The Local Musical Style of Kindergarten Children: a description and analysis of its natural variables. *Music Education Research*, 4, 1, 37-49.
- Gomes-Pedro, J. (1985). O comportamento do recém-nascido (2): os processos sensoriais. *Jornal de Psicologia*, 4, 3, 8-17.
- Gordon, E. E. (1990). *A Music Learning Theory for Newborn and Young Children*. Chicago: G.I.A. Publications.
- Gordon, E. E. (2000). *Teoria de aprendizagem musical para recém-nascidos e crianças em idade pré-escolar* (P. M. Rodrigues, Trans.). Lisboa: Fundação Calouste Gulbenkian.
- Guimarães, I. (2007). *A ciência e a arte da voz humana*. Alcabideche: Escola Superior de Saúde de Alcoitão.
- Henrique, L. (2002). *Acústica Musical*. Lisboa: Fundação Calouste Gulbenkian.
- Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100, 425-2438.
- Mang, E. (2001). A Cross-language Comparison of Preschool Children's Vocal Fundamental Frequency in Speech and Song Production. *Research Studies in Music Education*, 16, 4-14.
- Mertens, P. (2004). The Prosogram: Semi-automatic transcription of prosody based on a tonal perception model. In B. Bel & I. Marlien (Eds.) *Proceedings of Speech Prosody 2004*, Nara, Japan, pp. 23-26.
- Moog, H. (1976). The development of musical experience in children of pre-school age. *Psychology of Music*, 4, 38-45.
- Murray, L. (1991). Intersubjectivity, Object Relations Theory, and Empirical Evidence from Mother-Infant Interactions. *Infant Mental Health Journal*, 12 (3), 219-232.
- Papoušek, H. (1995). No princípio é uma palavra – uma palavra melodiosa. In J. Gomes-Pedro & M. F. Patrício (Eds.), *Bebé XXI – Criança e família na viragem do século* (pp. 171-175). Lisboa: Fundação Calouste Gulbenkian.
- Papoušek, H. (1996). Musicality in infancy research: biological and cultural origins of early musicality. In I. Deliège & J. Sloboda (Eds.), *Musical Beginnings: Origins and Development of Musical Competence* (pp. 37-51). Oxford: U. Press.
- Papoušek, M. (1996). Intuitive parenting. In I. Deliège & J. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 88-112). Oxford: University Press.
- Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences*, 1060, 59-70.
- Patel, A. D. (2008). *Music, Language, and the Brain*. New York: Oxford University Press.
- Reigado, J. (2007). Análise acústica das vocalizações de bebés de 9 a 11 meses face a estímulos musicais e linguísticos. *Dissertação de mestrado em Ciências Musicais. Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa*.
- Reigado, J. (2013). Desenvolvimento vocal na infância: análise acústica de vocalizações de bebés face a estímulos musicais e linguísticos durante o segundo ano de vida. *Dissertação de doutoramento em Ciências Musicais. Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa*.
- Reigado, J., Rocha, A., & Rodrigues, H. (2011). Vocalizations of infants (9-11 months old) in response to musical and linguistic stimuli. *International Journal of Music Education*, 29, 3, 241-255.
- Rivera, D. P., & Smith, D. D. (1997). *Teaching students with learning and behavior problems* (3rd ed.). Boston: Allyn and Bacon.
- Rocha, A. (2007). As vocalizações de bebés de 9 a 11 meses face à música e à linguagem. Análise efectuada por juízes especializados. *Dissertação de mestrado em Ciências Musicais. Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa*.
- Rodrigues, H. (2004, 14 de Abril). Desescolarizar a Educação. *Jornal de Letras*, p. 8:9.
- Rodrigues, H. (2005). A Festa da Música na iniciação à vida: da musicalidade das primeiras interações humanas às canções de embalar. *Revista da Faculdade de Ciências Sociais e Humanas*, 17, 61-80.
- Saffran, J. R., Loman, M. M., Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, 77, 15-23.
- Sacks, O. (2007). *Musicophilia*. New York: Alfred A. Knopf.
- Scotto Di Carlo, N. (2005). Contraintes de production et intelligibilité de la voix chantée. *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 24, 159-179.
- Scotto Di Carlo, N., & Autesserre, D. (1992). L'organisation temporelle de la syllabe dans la parole et dans le chant. *Travaux de l'Institut de Phonétique d'Aix*, 14, 189-204.
- Sundberg, J. (1987). *The Science of the Singing Voice*. DeKalb: Northern Illinois University Press.
- Sundberg, J. (1999). The perception of singing. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 171-214). San Diego: Academic Press.
- Tafari, J., & Villa, D. (2002). Musical elements in the vocalizations of infants aged 2-8 months. *British Journal of Music Education*, 19(1), 73-88.
- Trehub, S. (2003). Musical predispositions in infancy: an update. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 3-20). Oxford: University Press.

- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, M., & Saraza, M. (1997). Mothers and fathers' singing to infants. *Developmental Psychology*, 33, 500-507.
- Trevarthen, C. (1990). Early Parent-Child Interaction. In Th. A. Sebeok & J. Umiker-Sebeok, (Eds.), *The Semiotic Web* (pp. 41-51). New York: Mouton de Gruyter.
- Van Puyvelde M., Vanfleteren P., Loots G., Deschuyffeleer S., Vinck B., Jacquet W. & Verhelst W. (2010). Tonal Synchrony in Mother-Infant Interaction based on Harmonic and Pentatonic Series. *Infant Behavior and Development*, 4, 33, 387-400.
- Wallin, N., Merker, B. & Brown, S. (Eds.) (2000). *The origins of music*. Cambridge: MIT Press.
- Welch, G. (1994). The assessment of singing. *Psychology of Music*, 22, 3-19.
- Welch, G. F. (2006). Singing and Vocal Development. In G. McPherson (Ed.), *The Child as Musician: A handbook of musical development* (pp. 311-329). Oxford: Oxford University Press.
- Xavier, M. F., & Mateus, M. H. (Orgs.) (1990). *Dicionário de Termos Linguísticos*, Volume I. Lisboa: Edições Cosmos.
- Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, 111, 1399-1413.

Glossário

- análise segmental** – Descrição da estrutura geral de cada produção vocal da criança, isto é, da duração, do número de elementos constituintes e da velocidade de articulação vocal.
- durfonal** – Duração total de uma vocalização, descontando os tempos de silêncio entre fonações.
- durvocal** – Duração total de uma vocalização, correspondendo à soma das durações de fonação dentro e fora do núcleo com os tempos de silêncio entre fonações.
- frequência fundamental (F0)** – Frequência de vibração das cordas vocais, sendo um parâmetro fundamental na caracterização melódica da voz. Determina objetivamente a sensação de altura dos elementos de uma vocalização, permitindo ordenar os sons do grave ao agudo.
- interabsF0** – Soma, em valor absoluto, das variações da frequência fundamental (F0) entre núcleos sucessivos de uma vocalização.
- intradinF0** – Percentagem do tempo em que a variação de altura num núcleo excede o limiar de glissando. Ou seja, é a proporção da duração do núcleo em que a altura do som é entendida como movimento.
- limiar de glissando - G** – trata-se de um limiar auditivo para a percepção da variação da altura. Depende da amplitude (extensão) e da duração da variação frequência da fundamental (F_0). É normalmente expressa em ST/s (semitons

por segundo). O valor de referência para este estudo é $G = 0,32/T^2$, em que T é a duração da variação (cf. Mertens, 2004).

núcleo – Parte obrigatória da sílaba. Os núcleos das sílabas são sempre ocupados por vogais. Por exemplo: [i] corresponde ao núcleo inicial da palavra [bibe].

sílaba (sil) – Unidade de pronúncia maior que um simples som e menor que uma palavra (Xavier & Mateus, 1990). Cada sílaba obedece a uma organização hierárquica, sendo constituída pelo Ataque e a Rima. O Ataque corresponde à consoante ou consoantes com que se inicia a sílaba e a Rima integra um Núcleo (uma vogal ou ditongo) e pode conter uma Coda (uma consoante final).