



The CLEAN PAPER

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

Os mandarins tratam as vocalizações como categorias com significado — mas isso não é o mesmo que linguagem

Um estudo na Science testou se os mandarins apenas ouvem vocalizações diferentes ou se organizam o próprio repertório em categorias com significado comportamental. Em tarefas laboratoriais de discriminação, as aves aprenderam os 11 tipos de vocalização, generalizaram categorias a novos indivíduos e cometeram erros sistemáticos entre vocalizações usadas em contextos semelhantes mais frequentemente do que a semelhança acústica, por si só, previa. É evidência forte de percepção categorial e de uma forma operacional de significado. Não é evidência de que os mandarins tenham linguagem humana, sintaxe, palavras abertas ou um canal de conversa com pessoas.

Written by Laura Nesso · Cross-checked by Lucio Vaglio · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *Categorical and semantic perception of the meaning of call-types in zebra finches*, Julie E. Elie, Aude de Witasse-Thézy, Logan Thomas, Ben Malit, and Frédéric E. Theunissen, *Science* 389, 1210-1215 (2025) · DOI: 10.1126/science.ads8482

A palavra «significado» está a fazer aqui um trabalho cuidadoso

Os mandarins não precisam da nossa autorização para terem vida social. Chamam quando têm fome, estão separados, fazem corte, dão alarme, mantêm contacto com um parceiro ou entram em conflito. Os etólogos conseguem ordenar esses sons em **tipos de vocalização**: categorias práticas como chamadas de alarme, de contacto ou de pedido de alimento, definidas pela forma do som e pelo que a ave está a fazer quando o produz. A pergunta mais difícil é saber se as próprias aves organizam as vocalizações dessa maneira.

Um artigo na *Science* faz uma afirmação estreita, mas importante: mandarins adultos conseguem categorizar os tipos de vocalização do próprio repertório, e os seus erros sugerem que vocalizações associadas a comportamentos semelhantes ficam mais próximas no espaço percetivo das aves do que a acústica, por si só, permitiria prever.

É aqui que a palavra **semântico** entra no artigo. Não significa que as aves tenham palavras no sentido humano. Não significa sintaxe, conversa ou um dicionário escondido no cérebro de um mandarim. Aqui, «significado» é operacional: se duas vocalizações são usadas em contextos sociais semelhantes e as aves as confundem mais do que as propriedades sonoras, por si só, explicariam, então a categoria comportamental parece estar a moldar a percepção.

A versão limpa não é «as aves têm linguagem». É: os mandarins parecem tratar as suas próprias vocalizações como categorias funcionais, e algumas dessas categorias comportam-se como se transportassem significado no sentido limitado e testável da experiência.



Os mandarins (*Taeniopygia guttata*) são aves canoras altamente sociais, com um repertório rico de vocalizações, o que os torna um bom modelo para estudar como os animais categorizam sinais vocais.

Christoph Moning / iNaturalist CC BY 4.0

Meaning here is operational

Call categories are tested by behaviour, not by reading a bird's mind.

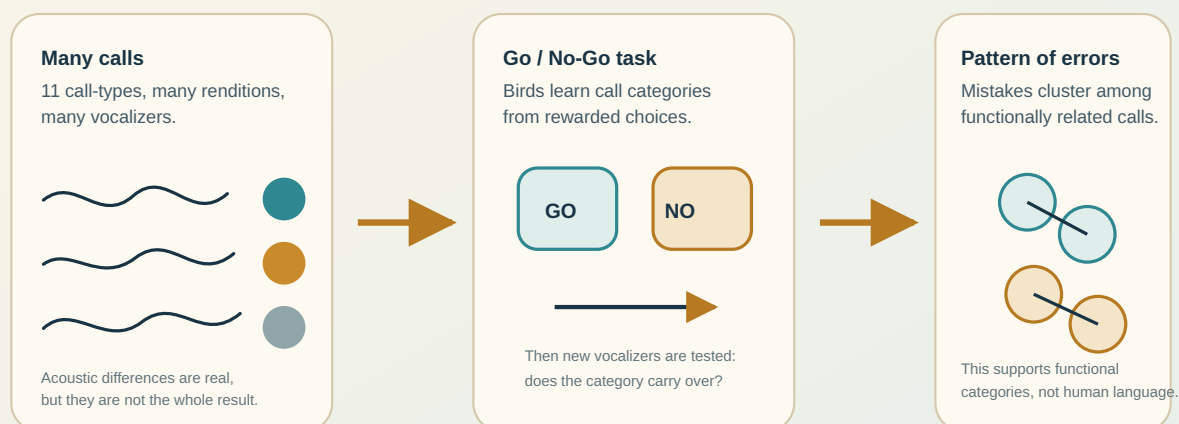


Diagrama preliminar da cadeia de evidência para revisão: muitas versões dos 11 tipos de vocalização entram numa tarefa de discriminação Go/No-Go; o resultado principal é que as aves generalizam as categorias a novos indivíduos e cometem erros agrupados pela função das vocalizações, não apenas pela semelhança acústica.

The Clean Paper CC BY 4.0

O que os autores testaram

Os autores partiram de um mapa etológico já existente do repertório do mandarim. A espécie tem cerca de **11 tipos de vocalização definidos por etograma**: vocalizações associadas a contacto, vínculo de casal e comportamento no ninho, alarme, agressão ou comportamento não afiliativo, pedido de alimento e canto dos machos.

Essa classificação foi feita por humanos. Combina características acústicas, o contexto em que um som é produzido e aquilo que o som tende a fazer na vida social. Mas um etograma humano não é automaticamente o mapa mental do animal. O estudo coloca três perguntas ligadas.

Primeiro, conseguem os mandarins discriminar todos esses tipos de vocalização?

Segundo, generalizam a categoria de um tipo de vocalização para indivíduos que nunca ouviram antes, em vez de simplesmente memorizar exemplos concretos?

Terceiro, quando cometem erros, seguem esses erros apenas a semelhança acústica ou também o significado comportamental das vocalizações?

O ponto crucial é que a experiência não pede a uma ave que explique o que uma vocalização significa.

Pergunta se o comportamento da ave numa tarefa controlada contém a assinatura estatística de categorias e significado.

Como funciona a experiência

Três termos têm muito peso neste artigo.

Um **tipo de vocalização** é uma categoria de som do mandarim. O artigo usa tipos de vocalização **definidos por etograma**, isto é, categorias herdadas de um catálogo comportamental: como o som se apresenta acusticamente, quando as aves o produzem e a situação social a que pertence. Um **vocalizador** é simplesmente a ave individual que produziu a gravação. Uma **forma acústica** é o padrão sonoro mensurável; uma **função comportamental** é aquilo para que a vocalização é normalmente usada.

Os 11 tipos de vocalização testados no artigo não são rótulos abstratos. São o repertório que os autores pediram às aves para discriminar:

- **Vocalizações de contacto:** distance call (DC), Tet (Te), long-tonal call (LT).
- **Vínculo de casal e comportamento no ninho:** whine call (Wh), nest call (Ne).
- **Vocalizações de alarme:** Tuck (Tu), Thuk (Th).
- **Vocalizações não afiliativas:** distress call (Di), aggressive call (Ag).
- **Pedido de alimento:** begging call (Be).
- **Corte:** canto do macho (So).

The 11 call-types are concrete categories

The paper asks whether zebra finches sort these calls the way ethologists do.

Contact calls

- DC distance call
- Te Tet
- LT long-tonal call

Pair-bonding / nest

- Wh whine call
- Ne nest call

Alarm calls

- Tu Tuck
- Th Thuk

Non-affiliative

- Di distress call
- Ag aggressive call

Begging

- Be begging call

Courtship

- So male song

These are ethogram-based labels: categories from sound shape plus social context, not human words.

Os 11 tipos de vocalização testados no artigo, agrupados por função social ampla. Os rótulos são categorias etológicas do artigo-fonte: não são «palavras», mas mostram ao leitor o repertório concreto que as aves tiveram de discriminar.

The Clean Paper CC BY 4.0

A tarefa principal foi uma **discriminação auditiva Go/No-Go**. Uma ave bicava uma tecla iluminada para iniciar uma reprodução de seis segundos. Perante uma vocalização recompensada, a resposta correta era esperar: depois de terminar o som, o comedouro subia e a ave recebia sementes. Perante uma vocalização não recompensada, esperar não dava nada; a resposta eficiente era voltar a bicar durante a reprodução, interromper o som e iniciar outro ensaio. Assim, as interrupções da ave revelam que sons ela trata como «não pertencentes à categoria recompensada».

Em cada teste, um tipo de vocalização era recompensado e os outros dez não. Os autores mudavam depois o tipo recompensado, percorrendo todo o repertório. Usaram também muitas versões produzidas por muitos vocalizadores, com poucas repetições exatas. Isto importa: a ave não podia simplesmente memorizar uma gravação. Para ter bom desempenho, tinha de aprender o que contava como aquele tipo de vocalização entre diferentes vozes e exemplos.

O posterior «espaço perceptivo» não é uma imagem cerebral nem um mapa literal dentro do animal. É um mapa inferido a partir dos erros. Se dois tipos de vocalização são frequentemente confundidos, são colocados mais próximos nesse mapa comportamental. Os autores compararam-no depois com um mapa acústico construído apenas a partir das características sonoras. A afirmação só se torna interessante se o mapa de erros das aves for puxado na direção da função da vocalização, e não apenas da semelhança sonora.

O que encontraram

As aves conseguiam discriminar todo o repertório. Doze mandarins adultos, seis machos e seis fêmeas, foram testados nos 11 tipos de vocalização. Quase todos os testes tiveram sucesso: **127 de 131** testes de discriminação foram significativos. As poucas falhas envolveram aves e tipos de vocalização específicos; o padrão global foi que todos os tipos foram discriminados acima do acaso.

Isto importa porque o repertório não é um conjunto arrumado de sinais sonoros isolados. Alguns tipos agrupam-se acusticamente, enquanto outros se transformam gradualmente uns nos outros. Ainda assim, as aves aprenderam as categorias.

Generalizaram as categorias a novos vocalizadores. Numa segunda experiência, as aves aprenderam a discriminar dois tipos de vocalização produzidos por um pequeno conjunto de indivíduos. No dia seguinte, foram acrescentados novos vocalizadores com a mesma regra de recompensa. As aves classificaram corretamente essas novas vocalizações logo no início do teste, antes de poderem aprender cada novo exemplo através do feedback da recompensa. Isto apoia percepção categorial do tipo de vocalização, e não apenas memorização de sons individuais.

A categoria conseguia sobrepor-se a uma nova regra de recompensa. O truque comportamental mais forte apareceu quando os autores tornaram a tarefa incongruente. Vocalizações de novos in-

divíduos receberam a contingência de recompensa oposta à que as aves tinham aprendido para aquele tipo de vocalização. Se as aves simplesmente seguissem os novos exemplos acústicos, deveriam adaptar-se imediatamente. Em vez disso, as respostas iniciais seguiram a categoria de tipo de vocalização já aprendida e produziram decisões de recompensa sistematicamente erradas. Ao longo do dia, as aves foram aprendendo lentamente a nova regra arbitrária. É exatamente o que se esperaria se a categoria natural do tipo de vocalização fosse suficientemente real para interferir.

Os erros não eram apenas acústicos. Os autores compararam um espaço acústico, construído a partir das confusões de um classificador entre sons, com um espaço perceptivo, construído a partir dos erros comportamentais das aves. Os dois espaços estavam relacionados: o som continuava a importar. Mas tipos de vocalização pertencentes à mesma hiper-categoria comportamental ou semântica ficavam mais próximos no espaço perceptivo das aves do que no espaço acústico. O artigo chama a isto **efeito de íman semântico**.

Em números, o agrupamento por hiper-categoria semântica era cerca de **2,43 vezes mais forte** no mapa perceptivo do que no mapa acústico. Em linguagem simples: os erros das aves eram puxados pela função, e não apenas por semelhanças sonoras.

O que «semântico» significa aqui

Esta é a parte que exige mais cuidado. Na linguagem quotidiana, «semântico» transforma-se rapidamente em «palavras». Não é isso que o artigo prova.

O estudo usa «semântico» num sentido comportamental e de cognição comparada. Um tipo de vocalização tem um significado proposto porque se associa a uma função social: alarme, contacto, pedido de alimento, agressão, vínculo de casal, corte. Se as aves confundem dois tipos da mesma categoria social ampla mais frequentemente do que a acústica prevê, então o seu espaço perceptivo está a ser moldado por essa categoria funcional.

É um resultado sério. Sugere que a ave não está simplesmente a ouvir um espectrograma. Está a tratar vocalizações específicas da sua espécie como categorias organizadas pelo comportamento.

Mas continua a ser indireto. Os autores não conseguem ler a mente do animal e dizem-no. Inferem representação interna a partir do comportamento: discriminação, generalização e erros sistemáticos.

Uma analogia útil não é «palavras de mandarim». É algo mais parecido com isto: o sistema auditivo da ave parece deformar a distância entre vocalizações de acordo com aquilo para que servem.

O que isto *não* prova

- **Não** mostra linguagem semelhante à humana. Não há sintaxe, gramática, significado composicional ou vocabulário aberto neste resultado.

- **Não** mostra que os mandarins tenham «palavras». Os tipos de vocalização são categorias vocais específicas da espécie ligadas a contextos comportamentais, não símbolos que possam ser recombinados livremente.
- **Não** mostra conversa com humanos nem um canal decodificado para falar com aves.
- **Não** acede diretamente a estados mentais. O significado é inferido a partir do comportamento na tarefa e da estrutura dos erros, não observado dentro da mente da ave.
- **Não** mostra que as aves compreenderam novos significados. A generalização a novos vocalizadores é generalização de uma categoria de vocalização através dos sons produzidos por diferentes indivíduos.
- **Não** resolve como estas categorias se desenvolvem. O estudo testou aves adultas; o papel da exposição e do desenvolvimento no efeito de íman semântico fica para trabalhos futuros.

Quão forte é a evidência?

Para a percepção categorial dos tipos de vocalização, a evidência é forte. O desenho pede às aves que discriminem todo o repertório, usa muitas versões de muitos indivíduos e inclui um teste de generalização em que vocalizações de indivíduos nunca ouvidos são classificadas por categoria. A condição de recompensa incongruente é especialmente útil porque mostra que a categoria consegue interferir com uma nova regra arbitrária.

Para um tipo operacional de percepção semântica, a evidência é boa, mas mais inferencial. O efeito de íman semântico não é apenas uma metáfora: os autores comparam mapas de distância acústica e comportamental e verificam que tipos de vocalização funcionalmente relacionados se agrupam mais fortemente na percepção do que na acústica. Isso apoia a ideia de que a função da vocalização molda a percepção.

Para significado semelhante ao humano, a evidência não existe — nem precisa de existir. O artigo é mais interessante se o deixarmos ser específico. A comunicação animal não ganha valor apenas quando se parece com a fala humana.

Porque é importante

A tentação pública é óbvia. Se uma ave ouve uma vocalização como significativa, a manchete seguinte quer dizer que estamos a decodificar a linguagem animal. Esse salto ignora a parte difícil da ciência.

O resultado cuidadoso é melhor. Mostra como investigadores podem testar «significado» sem fingirem traduzir a mente de um animal. Podem perguntar se os animais concordam com as categorias dos etólogos humanos. Podem perguntar se novos exemplos são ordenados pela categoria. Podem perguntar se os erros seguem semelhanças acústicas ou função social. É uma forma de tornar experimentalmente mais precisa uma pergunta antiga — as vocalizações animais significam alguma coisa para os próprios animais?

Também afasta a discussão de uma falsa escada em que os humanos têm linguagem e tudo o resto tem ruído. Os mandarins têm um repertório vocal pequeno e específico da espécie. Dentro desse repertório, este estudo sugere que as aves percebem categorias estruturadas ligadas ao uso social. Não é a nossa linguagem. É o sistema de comunicação delas, testado nos próprios termos.

Resumo claro

Um estudo na *Science* testou mandarins adultos nos 11 tipos de vocalização do seu repertório. Em tarefas auditivas Go/No-Go, as aves discriminaram todos os tipos, generalizaram categorias aprendidas para vocalizações de novos indivíduos e, numa condição incongruente, seguiram inicialmente as categorias naturais mesmo quando isso produzia decisões de recompensa erradas. Os autores compararam depois semelhança acústica com erros comportamentais e verificaram que vocalizações usadas em contextos sociais semelhantes se agrupavam mais fortemente no espaço perceptivo das aves do que no espaço acústico. Isto sustenta percepção categorial e uma forma operacional de percepção semântica: as aves parecem organizar as próprias vocalizações por categorias funcionais, não apenas pelo som. Não mostra linguagem humana, sintaxe, palavras, acesso direto a estados mentais nem conversa com aves.

Verificação sem exageros

O que o artigo mostra: Mandarins adultos conseguem discriminar os 11 tipos de vocalização definidos por etograma do seu repertório; generalizam as categorias para novos vocalizadores; e os seus erros sistemáticos são moldados por categorias comportamentais ou semânticas para além da semelhança acústica.

O que é plausível, mas não está provado: Que os mandarins tenham representações internas do significado das vocalizações; que mapas neurais semelhantes estejam por trás do «efeito de íman semântico» comportamental; que desenvolvimento e exposição moldem estas categorias de forma análoga a outras categorias perceptivas aprendidas.

O que não mostra: Linguagem semelhante à humana; gramática; palavras abertas; referência simbólica no sentido humano; evidência direta de estados mentais subjetivos; uma forma de humanos falarem com mandarins.

Principais limitações: O trabalho usa tarefas operantes em laboratório com **12 aves adultas**, não interações naturais no campo; «semântico» é inferido a partir do comportamento e da estrutura dos erros; o desenvolvimento continua em aberto; o resultado diz respeito a um repertório fixo e específico da espécie, não a linguagem composicional flexível.

Quanta confiança deve ter um leitor geral? Confiança elevada de que os mandarins conseguem categorizar e discriminar os seus tipos de vocalização nesta tarefa. Boa confiança de que os erros refletem categorias funcionais, não apenas semelhança acústica. Confiança baixa de que isto consti-

tua evidência de linguagem de aves no sentido humano. A atitude adequada: um resultado forte em cognição animal sobre categorias vocais com significado funcional, não um momento Dr. Dolittle.

Fontes

Science 389, 1210-1215 (2025)

<https://doi.org/10.1126/science.ads8482>

Author manuscript in PMC — PMCID PMC12721838

<https://pmc.ncbi.nlm.nih.gov/articles/PMC12721838/>

Behavioral data and analysis code — Zenodo

<https://zenodo.org/records/13623969>

Zebra finch acoustical database — Figshare

<https://doi.org/10.6084/m9.figshare.11905533.v1>