



The CLEAN PAPER

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

斑胸草雀把叫声视为有意义的类别——但这并不等于语言

一项 *Science* 研究检验斑胸草雀究竟只是听见不同叫声，还是会把自已的鸣声库组织成承载行为意义的类别。在实验室辨别任务中，鸟学会了全部 11 种叫声类型，把类别泛化到新的发声个体，并且混淆用于相似情境的叫声，其频率高于单凭声学相似度所能预测的水平。这是类别知觉和一种操作性意义的有力证据，但不是斑胸草雀拥有类人语言、句法、开放式词语或与人对话渠道的证据。

Written by Laura Nesso · Cross-checked by Lucio Vaglio · Edited by Michele Renda · Figures by Laura Nesso
· 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

“意义”一词在这里用得很谨慎

斑胸草雀不需要得到我们的许可，才拥有自己的社会生活。它们会在饥饿、分离、求偶、警戒、与伴侣保持联系或发生冲突时鸣叫。动物行为学家可以把这些声音归入不同的**叫声类型**：例如警戒叫声、联络叫声或乞食叫声。这些实用类别由声音的形态，以及鸟发出声音时正在做什么来界定。更难的问题是，鸟自己是否也用这种方式给叫声分类。

一篇发表于 *Science* 的论文提出了一个范围狭窄但重要的主张：成年斑胸草雀能够区分自身鸣声库中的各种叫声类型；而它们所犯的错误表明，在鸟的知觉空间中，行为用途相关的叫声彼此更接近，超过了单凭声学特征所能预测的程度。

这正是论文引入**语义**一词的地方。它并不意味着这些鸟拥有人类意义上的词语，也不意味着语法、对话，或雀脑里藏着一本词典。这里的“意义”是操作性的：如果两种叫声用于相似的社会情境，而鸟混淆它们的频率高于声音本身所能解释的程度，那么行为类别似乎正在塑造知觉。

清晰的说法不是“鸟类拥有语言”，而是：斑胸草雀似乎把自己的叫声当作功能类别；其中一些类别在实验限定、可以检验的意义上，表现得仿佛承载着意义。



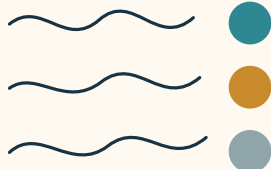
斑胸草雀 (*Taeniopygia guttata*) 是高度社会化的鸣禽，拥有丰富的叫声库，因此很适合用来研究动物如何对声音信号分类。

Meaning here is operational

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

Many calls

11 call-types, many renditions,
many vocalizers.



Acoustic differences are real,
but they are not the whole result.

Go / No-Go task

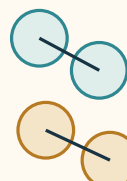
Birds learn call categories
from rewarded choices.



Then new vocalizers are tested:
does the category carry over?

Pattern of errors

Mistakes cluster among
functionally related calls.



This supports functional
categories, not human language.

供审阅的证据链幻灯片草稿：11 种叫声类型的大量实例进入一项 Go/No-Go 辨别任务；关键结果是，鸟能够把类别泛化到新的发声个体，而且错误会按叫声功能聚集，而不只按声学相似度聚集。

The Clean Paper CC BY 4.0

作者检验了什么

作者从现有的斑胸草雀鸣声行为图谱出发。这个物种大约有 **11 种基于行为谱的叫声类型**：分别与联络、配偶关系和筑巢行为、警戒、攻击或非亲和行为、乞食以及雄性鸣唱相关。

这个分类由人类完成，综合了声学特征、声音产生时的情境，以及它在社会生活中通常起什么作用。但人类的行为谱不会自动等同于动物的心理地图。这项研究提出了三个彼此关联的问题。

第一，斑胸草雀能否辨别所有这些叫声类型？

第二，它们能否把某种叫声类型泛化到此前没有听过的发声个体，而不只是记住单个样本？

第三，当它们出错时，错误是否只遵循声学相似度，还是也会遵循叫声的行为意义？

关键在于，实验并不要求鸟解释某种叫声是什么意思。它问的是：鸟在受控任务中的行为，是否留下了类别和意义的统计印记。

实验怎样进行

这篇论文中有三个术语承担了很多含义。

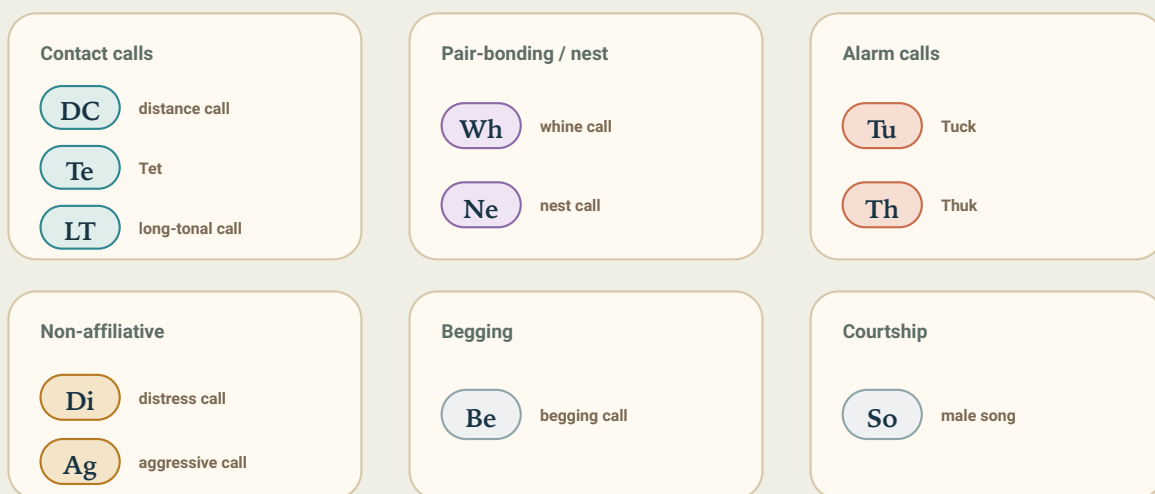
叫声类型是斑胸草雀声音的一种类别。论文采用**基于行为谱**的叫声类型，也就是继承自行为目录的类别：声音在声学上是什么样，鸟在什么时候发出它，以及它属于哪种社会情境。**发声个体**就是产生某段录音的那只鸟。**声学形态**是可以测量的声音模式；**行为功能**则是这种叫声通常用来做什么。

论文检验的 11 种叫声类型不是抽象标签，而是作者要求鸟辨别的实际鸣声库：

- **联络叫声**：远距离叫声（DC）、Tet 叫声（Te）、长音调叫声（LT）。
- **配偶关系和筑巢行为**：呜咽叫声（Wh）、巢叫声（Ne）。
- **警戒叫声**：Tuck 叫声（Tu）、Thuk 叫声（Th）。
- **非亲和叫声**：痛苦叫声（Di）、攻击叫声（Ag）。
- **乞食**：乞食叫声（Be）。
- **求偶**：雄性鸣唱（So）。

The 11 call-types are concrete categories

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



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

论文检验的 11 种叫声类型，按宽泛的社会功能分组。这些标签是来源论文中基于行为谱的类别：它们不是“词语”，但让读者看到鸟必须辨别的具体鸣声库。

The Clean Paper CC BY 4.0

主要任务是一项 **Go/No-Go 听觉辨别**。鸟啄一下亮起的按键，开始播放一段六秒声音。如果是有奖励的叫声，正确做法是等待：声音结束后，食槽升起，鸟得到种子。如果是无奖励叫声，等待不会得到任何东西；更有效的做法是在播放期间再次啄键，中断声音并开始下一次试验。因此，鸟何时打断播放，显示了它把哪些声音视为“不属于有奖励类别”。

每轮测试以一种叫声类型为有奖励类别，其余十种没有奖励。随后作者更换有奖励的叫声类型，遍历整个鸣声库。他们还使用了许多发声个体的大量不同实例，几乎不精确重复。这一点很重要：鸟不能只记住某一段录音。要表现良好，它必须学会跨越不同鸟的声音和不同样本，判断什么才算属于该叫声类型。

后文所说的“知觉空间”不是脑扫描，也不是动物脑中的一张字面地图。它是一张由错误推断出来的地图。如果两种叫声类型经常被混淆，它们在这张行为地图中的位置就更接近。作者把它与仅由声音特征构成的声学地图进行比较。只有当鸟的错误地图被叫声功能拉动，而不只是被声音相似度拉动时，这个主张才真正有意思。

他们发现了什么

鸟能够辨别完整的鸣声库。研究测试了 12 只成年斑胸草雀，雌雄各六只，覆盖 11 种叫声类型。几乎所有测试都成功：131 次叫声类型辨别测试中，**127 次**达到显著水平。少数失败涉及特定鸟和特定叫声类型；总体模式是，每种叫声类型的辨别表现都高于随机水平。

这一点很重要，因为这个鸣声库并不是一组彼此孤立、整齐划一的提示音。有些叫声类型在声学上聚集成簇，另一些则彼此渐变。鸟仍然学会了这些类别。

它们把类别泛化到了新的发声个体。在第二项实验中，鸟先从少量发声个体那里学习辨别两种叫声类型。第二天，研究者加入新的发声个体，并沿用同一奖励规则。鸟在测试初期就正确分类了这些新叫声，此时它们还来不及通过奖励反馈逐一学会所有新样本。这支持对叫声类型进行类别知觉的解释，而不只是记忆单个声音。

类别可以压过新的奖励规则。最有力的行为设计出现在作者让任务发生不一致时。新发声个体的叫声被赋予与鸟已经学会的该叫声类型相反的奖励条件。如果鸟只是跟随新的声学样本，就应该立刻适应。相反，它们在初期仍按照此前学会的叫声类型类别作答，因此系统性地做出错误的奖励判断。随着一天的进行，鸟才逐渐学会这条人为的新规则。如果天然叫声类型的类别真实存在，而且强到足以造成干扰，这正是应当看到的现象。

错误不只是声学造成的。作者比较了两种空间：一是由分类器混淆不同叫声的情况构成的声学空间，二是由鸟的行为错误构成的知觉空间。两者确实相关：声音仍然重要。但属于同一行为或语义超类别的叫声类型，在鸟的知觉空间中彼此更接近，而且这种接近程度超过声学空间。论文把这称为**语义磁铁效应**。

用数字表示，按语义超类别分组的强度，在知觉地图中约为声学地图的 **2.43 倍**。通俗地说：鸟的错误被功能意义拉近，而不只是被相似的声音拉近。

“语义”在这里是什么意思

这是最需要谨慎的部分。在日常语言中，“语义”很快就会被理解成“词语”。这不是论文证明的内容。

研究是在行为学和比较认知的意义上使用“语义”。一种叫声类型具有推定意义，是因为它与某种社会

功能相关：警戒、联络、乞食、攻击、配偶联系或求偶。如果鸟混淆同一宽泛社会类别中两种叫声的频率高于声学特征所能预测的水平，那么它们的知觉空间就受到了这个功能类别的塑造。

这是一个严肃的结果。它表明，鸟听到的并不只是一张声谱图，而是把本物种特有的叫声当作按行为组织的类别。

但证据仍是间接的。作者无法读取动物的心智，他们也明确承认这一点。内部表征是从行为中推断出来的：辨别、泛化和系统性错误。

一个有用的类比不是“雀类词语”，而更接近于：鸟的听觉系统似乎会根据叫声的用途，扭曲不同叫声之间的距离。

这项研究没有证明什么

- 它**没有**展示类人的语言。结果中没有句法、语法、组合意义或开放式词汇。
- 它**没有**展示斑胸草雀拥有“词语”。这些叫声类型是与行为情境相连的物种特有声音类别，不是可以自由重组的符号标签。
- 它**没有**展示与人类对话，也没有解码出一条让人和鸟交流的渠道。
- 它**没有**直接读取心理状态。意义是从任务行为和错误结构中推断出来的，不是从鸟的心智内部观察到的。
- 它**没有**展示鸟理解了新的意义。对新发声个体的泛化，是跨越不同个体声音对叫声类型类别的泛化。
- 它**没有**解决这些类别如何发育的问题。研究测试的是成年鸟；接触经验和发育如何塑造语义磁铁效应，仍是未来工作。

证据有多强？

对于叫声类型的类别知觉，证据很强。实验设计要求鸟辨别完整的鸣声库，使用许多发声个体的大量实例，并包含一项泛化测试，观察鸟是否按叫声类型分类此前未听过的发声个体。不一致奖励条件尤其有用，因为它表明类别能够干扰一条新的人为规则。

对于一种操作性语义知觉，证据良好，但推断性更强。语义磁铁效应不只是一个比喻：作者比较声学距离图和行为距离图，发现行为用途相关的叫声类型在知觉中比在声学上聚集得更紧密。这支持叫声功能会塑造知觉的观点。

至于类人的意义，证据并不存在，也没有必要存在。只要让论文保持具体，它反而更有意思。动物交流的价值，并不取决于它是否像人类语言。

为什么重要

公众很容易受到诱惑：如果一只鸟把某种叫声听作有意义，下一个标题就想说我们正在解码动物语言。这个跳跃绕过了科学中最艰难的部分。

谨慎的结果更好。它展示了研究人员如何检验“意义”，又不假装自己能翻译动物的心智。他们可以问，动物是否认同人类行为学家的分类；可以问，新样本是否按类别归类；还可以问，错误究竟遵循声学相似度，还是社会功能。这样就能把一个古老的问题——动物叫声对动物自己是否意味着什么？——变得更适合实验检验。

它也让讨论摆脱了一架虚假的阶梯：人类在顶端拥有语言，其他一切都只是噪声。斑胸草雀拥有一个规模不大、物种特有的鸣声库。在这个鸣声库中，研究表明它们知觉到与社会用途相连的结构化类别。这不是我们的语言，而是它们的交流系统，按照它自身的尺度接受检验。

清晰摘要

一项发表于 *Science* 的研究用斑胸草雀鸣声库中的 11 种叫声类型测试成年鸟。在听觉 Go/No-Go 任务中，鸟能辨别所有叫声类型，把学到的类别泛化到新发声个体；而在不一致条件下，即使按叫声类型作答会导致错误的奖励判断，它们起初仍会遵循叫声类型类别。随后，作者比较声学相似度和行为错误，发现用于相似社会情境的叫声类型在鸟的知觉空间中，比在声学空间中聚集得更紧密。这支持类别知觉和一种操作性语义知觉：鸟似乎按功能类别，而不只按声音，来组织自己的叫声。研究没有展示类人的语言、句法、词语、对心理状态的直接读取，也没有展示人与鸟的对话。

不绕弯核查

论文证明了什么：成年斑胸草雀能够辨别其鸣声库中 11 种基于行为谱的叫声类型；它们会把叫声类型类别泛化到新的发声个体；其系统性错误除声学相似度外，还受行为或语义类别塑造。

什么合理但尚未证明：斑胸草雀拥有叫声意义的内部表征；类似的神经地图支撑行为上的“语义磁铁”效应；发育和接触经验以类似其他习得知觉类别的方式塑造这些类别。

它没有证明什么：类人的语言；语法；开放式词语；人类意义上的符号指代；主观心理状态的直接证据；让人类与斑胸草雀交流的方法。

主要局限：研究采用实验室操作性任务，测试的是 **12 只成年鸟**，而非自然野外互动；“语义”从行为和错误结构推断；发育问题仍未解决；结果针对固定的物种特有鸣声库，不是灵活的组合式语言。

普通读者应该有多大信心？可以高度确信斑胸草雀能在这项任务中对叫声类型进行分类和辨别。可以较有信心地认为，它们的错误反映了功能类别，而不只反映声学相似度。几乎没有理由把它当成人类意义上的鸟类语言证据。恰当态度是：这是一项关于有意义声音类别的强有力动物认知结果，不是杜立德式的突破。

来源

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>