LSVT数据集，14名帕金森患者参与（都是PD患者进行采集），采集的语音用来训练并判别语音是否可以接收。

126个样本，每个样本有309个特征，其中可以接收的语音样本标签为1，有42个样本。

不可以接收的语音样本标签为2 ，有82个。

不可以接受的判断为 有病患者， 也就是标签2

可以接受的 判断为 康复患者，也就是标签1 。

所以lsvt数据集中 标签1 康复患者， 标签2 为 未康复患者

帕金森氏病康复性言语治疗的客观自动评估

摘要

语音性能下降是绝大多数帕金森氏病（PD）患者的常见症状，他们通常会长期与语音专家进行一对一的定期定期康复会议。 最近，开发了一种名为Lee Silverman语音治疗（LSVT）伴侣的新型计算机程序，以使PD受试者能够通过康复治疗过程独立进展。 这项研究是LSVT伴奏评估的一部分，旨在调查使用持续元音发声的可能性，以便客观，自动地将语音专家对PD对象声音的评估复制为“可以接受的”（临床医生会允许在持续进行期间 人康复治疗）或“不可接受”（临床医生不允许在人康复治疗期间坚持治疗）。 我们用309个声韵测度来表征156个持续元音/ a /发声中的每一个，使用稳健的特征选择算法选择一个简约的子集，并自动以大约90％的总体准确度区分这两个队列（可接受与不可接受）。 此外，我们说明了所提出的方法作为语音专家将概率评估为“可接受”或“不可接受”的发声的概率决策支持工具的潜力。 我们认为这项研究的结果是提高自动康复语音评估工具有效性的第一步。

摘要

对于绝大多数帕金森病（PD）受试者来说，嗓音表现退化是一种常见症状，他们通常会在长期内与言语专家进行个性化的一对一定期康复会议。最近，一个叫做Lee Silverman语音治疗（LSVT）Companion的新的计算机程序被开发出来，允许PD患者通过康复治疗独立地进步。这项研究是LSVT同伴评估的一部分，旨在研究使用持续元音发音客观地自动复制言语专家对帕金森病患者声音的评估为“可接受”（临床医生允许在康复治疗过程中坚持）或“不可接受”（临床医生不允许在治疗期间坚持）亲自康复治疗）。我们对156个持续元音/a/发音进行了309个发音困难测试，使用稳健的特征选择算法选择一个简约子集，并以大约90%的整体准确率自动区分两个队列（可接受与不可接受）。此外，我们说明了所提出的方法作为一个概率决策支持工具的潜力，语音专家评估一个发音是“可接受的”或“不可接受的”。我们设想这项研究的发现是提高自动康复语音评估工具有效性的第一步。

什么叫做可以接受，就是患者通过发音，以生觉得他已经达到康复的标准，

不可以接受则相反

1.引言

帕金森氏病（PD）是一种慢性神经退行性疾病，其特征是运动功能逐步恶化和出现大量非运动问题[1]。据估计，人口中每100000至少有100个PD受试者[2]，并且一些研究表明PD的患病率 可能会被低估[3]。 在绝大多数的PD受试者中都报告了声音障碍，并且大约29％的与该疾病相关的不良反应中考虑了吗啡酮[4]。 而且，语音表现下降可能是在PD发作的首发症状中[5]。

典型的声音障碍症状包括响度降低，单调，声音嘶哑，呼吸（噪声），不精确的发音和声音震颤[6]。 可以通过持续的元音发声或慢跑的语音来评估语音损伤的程度。 可以说，尽管持续性元音中可能不会发生由PD引起的跑步语音中的某些语音缺陷（例如，辅音和元音序列），但由于发音和其他语言上的混淆，对跑步语音的分析从本质上来说更加复杂[ 7]，[8]。 因此，在临床实践中通常会使用持续的元音发声，即说话人试图产生尽可能稳定（在振幅和频率方面）且尽可能长的元音。 临床实践和大量研究均表明，持续元音“ ahh”（表示为/ a /）可能足以用于许多语音评估应用[8] – [11]，尤其是用于PD语音评估[12] – [17] ]。

临床语音信号处理算法（从语音信号中提取临床有用信息的算法工具）被统称为声音障碍测量。 从语音专家和耳鼻喉外科医师的角度来看，语音（由喉部产生的声音）与语音之间存在重要的区别，语音和发音是语音的综合过程。 尽管在语音和语音的狭义定义上存在细微的差异，但I. Titze断言“在广义上，语音与语音同义” [8]。 严格来说，本研究中讨论的声音障碍测量完全基于声音而不是语音，因为我们没有尝试描述发音问题。

LSVT LOUD是一种基于研究的标准化语音治疗方案，对PD的疗效已确立[18] – [20]。一个重大挑战是如何扩展这种语音治疗计划的可及性，而如果语音处理计划不足，则需要持续的强化治疗方案 临床医生可以提供所需的所有治疗。 计算机和基于Web的技术的进步为治疗可及性，有效的治疗提供以及康复中的长期维护提供了解决方案[21]。 这样的技术可以减轻临床医生人数不足以提供亲自治疗的障碍，增强提供密集治疗要求的可行性，并减轻往返于诊所进行亲自治疗的后勤负担。 通过美国国立卫生研究院（NIH）和迈克尔·J·福克斯基金会（Michael J. Fox Foundation）的资助，LSVT Global开发并测试了LSVT Companion，这是一种计算机程序，可让PD受试者在治疗过程中独立进行。 和语音任务期间的发声持续时间，并在治疗练习期间向客户提供有关他们的响度，音调和发声持续时间的瞬时音频和视觉反馈。 当PD对象未达到其目标目标（声音响度，音调，持续时间）时，LSVT伴侣会提供反馈以鼓励他们适当调整其发声特性，例如更大声，更长，更高或更低。

LSVT LOUD是一种标准化的、基于研究的语音治疗方案，对PD[18]-[20]具有既定的疗效。一个重要的挑战是，当没有足够的临床医生提供所需的所有治疗时，如何扩大这种需要持续强化治疗方案的可及性。计算机和网络技术的进步为康复治疗的可及性、有效的治疗提供了解决方案，以及长期维护[21]。这种技术可以缓解临床医生人数不足的障碍，提高提供强化治疗要求的可能性，并减轻往返诊所进行亲自治疗的后勤负担。通过美国国立卫生研究院（NIH）和Michael J.Fox基金会的资助，LSVT Global已经开发并测试了LSVT Companion，这是一个允许帕金森病患者独立地通过ughatreatmentsession.创建会议LSV公司自动平衡声压级（SPL）、基频（F0），以及在语音任务期间的发声持续时间，并在治疗练习期间向客户提供关于其声音响度、音调和发声持续时间的即时视听反馈。当PD受试者没有达到他们的目标目标（声音的响度、音调、持续时间），LSVT同伴会提供反馈，鼓励他们适当地调整他们发声的特征，例如声音更大、更长、更高或更低。

在最近的一项研究中，LSVT伴侣被用于增加面对面的治疗疗程[22]。在16个疗程中，有7个疗程的受试者接受了独立的家庭治疗，其中9个疗程是传统的面对面治疗。LSVT后即刻和6个月后SPL的结果数据与16次临床医生亲自治疗的数据相比较。LSVT伴侣开发的下一步是进一步支持独立治疗（即增加独立的家庭治疗次数）。一个潜在的担忧是，在学习提高声音响度的过程中，帕金森病患者可能会表现出不可接受的声音特征，而LSVT专家临床医生在当面治疗期间“不允许持续存在”。

本研究的目的是研究使用一个客观的统计机器学习框架自动评估持续元音发音为“可接受”（临床医生允许坚持语音治疗）或“不可接受”（临床医生不允许坚持语音治疗）的可能性。最终的目标是通过为LSVT伴奏系统开发合适的算法来提高康复语音治疗的有效性。该算法能够在远离专家临床指导的情况下检测不可接受的语音特征，阻止患者以不可接受的方式使用语音，并通过反馈改善语音特征。

本研究的目的是研究使用一个客观的统计机器学习框架自动评估持续元音发音为“可接受”（临床医生允许坚持语音治疗）或“不可接受”（临床医生不允许坚持语音治疗）的可能性。最终的目标是通过为LSVT伴奏系统开发合适的算法来提高康复语音治疗的有效性。该算法能够在远离专家临床指导的情况下检测不可接受的语音特征，阻止患者以不可接受的方式使用语音，并通过反馈改善语音特征。

本研究的目的是研究使用一个客观的统计机器学习框架自动评估持续元音发音为“可接受”（临床医生允许坚持语音治疗）或“不可接受”（临床医生不允许坚持语音治疗）的可能性。最终的目标是通过为LSVT伴奏系统开发合适的算法来提高康复语音治疗的有效性。该算法能够在远离专家临床指导的情况下检测不可接受的语音特征，阻止患者以不可接受的方式使用语音，并通过反馈改善语音特征。

2 .数据

最初筛选出17名帕金森病患者纳入本研究，其中14名最终参与。三名没有参与这项研究的受试者是由于与数据收集会议的日程安排冲突。所有受试者均具有PD的典型语音和语音特征，由经验丰富的言语语言病理学家通过电话筛查确定（例如，响度降低、单调、呼吸困难、声音沙哑或发音不准确），并在数据收集过程中由两名具有PD专业知识的语言病理学家进行验证。受试者在这项研究中的注册和所有的招募材料都是由一个独立的机构审查委员会批准的。

14名帕金森病受试者（8名男性和6名女性）的年龄范围为51-69岁（平均标准差：61.9 6.5）并产生持续的元音/a/发音。这些持续的元音发音是在丹佛国家语音和语音中心（NCVS）的一个双壁、消音的房间里录制的，该中心是科罗拉多大学博尔德分校的附属机构。头戴式麦克风根据标准协议进行定位[23]。语音信号在44.1khz，16位分辨率下采样，并使用Audacity软件包进行记录。总的来说，每个受试者最初被要求产生27个发音（样本），其中每个发音属于音调和振幅的九种可能组合之一，即舒适音高、高音、低音，振幅可接受，太大或者太柔软。后面的理由使用不同的音高和振幅条件可以从不同类型的发声中获得样本；研究表明，有意识地思考我们说话的方式可能会提高声乐表现[24]。其目的是从九种组合中每一种都使用一个好的样本：在实践中，记录好的样本不仅仅是一次尝试（一些受试者大笑或咳嗽）。同样地，一些受试者需要不止一次的尝试来遵循要求他们完成的任务（例如，当刺激高振幅发声时）。然后，对于每一个受试者，数据收集小组（在任何语音信号分析之前）从九种可能的音调和幅度组合中选出最佳样本，并将每个样本评估为“可接受”或“不可接受”。这一数据选择步骤是为了确保我们为每个受试者处理相同数量的样本，对每个受试者的音高和振幅的每个组合使用一个良好的代表性样本，避免了可能不够“好”的声音污染数据集。丢弃的声音不符合目标标准（例如，对于“高振幅”任务，声音不够大，当被要求发出低沉或高音的声音、咳嗽或大笑等时，音高没有变化。因此，每个PD受试者共处理了126个发音，共有9个。在这126个样本中，有25%的样本被重复，以量化总共156个样本的内部评分信度。

LSVT专家临床医生通过感知来评估语音是否可以被认为是“可接受”还是“不可接受”。语音专家没有具体的客观标准来评估发音；评估很大程度上取决于评估者的经验。根据两位经验丰富的言语专家（福克斯和拉米格）的一致评估，每个受试者都产生了两种类型的发音（“可接受”和“不可接受”）。这项评估将被视为“基本事实”。我们还对每一个发音进行了五次额外的评估，这些专家对其他评分员的评估和受试者的状况一无所知。在所有情况下，专家评分员必须决定每个发音是“可接受”还是“不可接受”。因此，我们对156个样本进行了6次评估。

每个发声都经过预处理，从发音的稳定部分（语音信号的中间部分）中选择一个2s段进行后续声学分析，以避免在发声开始（开始）和偏移（结束）期间出现问题。

人工检查没有发现任何有问题的记录（除了“啊”的声音，例如咳嗽），所以在分析中使用了所有的数据样本。