[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G09B 19/04 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200780018936.5

[43] 公开日 2009 年 6 月 10 日

[11] 公开号 CN 101454816A

[22] 申请日 2007.5.11
[21] 申请号 200780018936.5
[30] 优先权
[32] 2006. 5.22 [33] EP [31] 06114304.6
[86] 国际申请 PCT/IB2007/051787 2007.5.11
[87] 国际公布 WO2007/135605 英 2007.11.29
[85] 进入国家阶段日期 2008.11.24
[71] 申请人 皇家飞利浦电子股份有限公司
地址 荷兰艾恩德霍芬
[72] 发明人 R·D·威尔曼 G·兰弗曼
D·格勒
[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 李亚非 刘 红

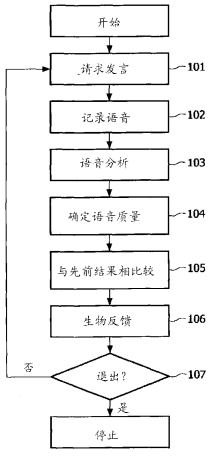
权利要求书 1 页 说明书 8 页 附图 2 页

[54] 发明名称

训练构音障碍的说话者的系统和方法

[57] 摘要

本发明描述了一种基于计算机的系统，其中该系统要求(101)病人发出显示在监视器上的单词的读音，自动评定(104，105)语音质量，并且使用适当的装置来反馈(106)语音质量改善或恶化。



1. 一种用于训练构音障碍的说话者的系统（1），包括:

第一计算模块(4),适于自动确定说话者的输入话语的语音质量量度，

第二计算模块（5），适于自动将语音质量量度与说话者至少一个先前的语音质量量度相比较，以及

反馈模块（6），适于根据该比较来自动向说话者提供生物反馈。

2. 如权利要求 1 所述的系统（1），包括: 记录设备（3），适于记录说话者的输入话语。

3. 如权利要求 1 所述的系统（1），包括: 接口模块（6），适于向说话者提供说话指令。

4. 如权利要求 1 所述的系统（1），其中，第一计算模块（4）包括语音分析组件（11），适于在音位和/或单词级别上执行语音分析。

5. 如权利要求 1 所述的系统（1），其中，第一计算模块（4）包括比较组件（12），适于将输入话语与参考数据相比较，所述参考数据是以多个非构音障碍的说话者的数据为基础的。

6. 如权利要求 1 所述的系统（1），其中，第二计算模块（5）包括评估组件（13），适于相对于说话者的先前结果执行语音质量量度的比较评估。

7. 一种训练构音障碍的说话者的方法，包括以下步骤:

自动确定（104）说话者的输入话语的语音质量量度，

自动将语音质量量度与说话者的至少一个先前语音质量量度相比较（105），以及

根据该比较，自动向说话者提供（106）生物反馈。

8. 一种在计算机（2）中执行的计算机程序，当在计算机（2）中执行该计算机程序时，所述程序包括:

用于自动确定（104）说话者的输入话语的语音质量量度的计算机指令，

用于自动将该语音质量量度与说话者的至少一个先前语音质量量度相比较（105）的计算机指令，以及

用于根据所述比较来自动向说话者提供（106）生物反馈的计算机指令。

训练构音障碍的说话者的系统和方法

本发明涉及一种训练构音障碍的说话者(dysarthric speaker)的系统和方法。此外，本发明还涉及一种用于训练构音障碍的说话者的计算机程序。

语音障碍(speech deficit)是中风之类的神经系统疾病的一个共同症状。在美国，每年都会新发 700,000 起中风。虽然在这些病例中只有20%是致命的，但是中风将会导致 60%的受害者受到损伤。大约 25%的病人将会显现出语音障碍。每年，康复措施和劳动力损失会使美国社会每年承受 500 亿美元的代价。

目前有两大组语音障碍：失语症和构音障碍。失语症是指病人没有能力正确叫出物体的名称。而构音障碍则是指病人没有能力清楚、正确地发音。

语音治疗学家通常会通过让病人重复某个单词，以便在一对一的会话中训练构音障碍的说话者。而病人则并未接收到详细的纠正。很明显，这种治疗方式的成本非常密集，并且对治疗学家的要求很高。

为了在优化成本的同时将康复成果增至最大，较为理想的是由病人在没有监理专业人员可用的情况下在家中继续进行康复练习。众所周知，恰当的反馈将会增强康复成功率。为此目的，有必要自动对练习进行评定，例如借助计算机来进行评定。

在计算机上，失语症的测试相对容易，例如，该测试可以通过显示物体的图像以及通过以多选方式向病人询问该物体的正确名称来进行。当前，训练软件业已得到商用，并且在康复实践中得到了使用。但是，由于语音状况有可能会很严重，因此，在计算机上是很难测试构音障碍的。当前，用于以无人监管的方式训练受损语音的系统只能用于失语症，而不能用于构音障碍。

本发明的一个目的是提供一种训练构音障碍的说话者的廉价、可靠且无人监管的技术、

这个目标根据本发明是通过一种用于训练构音障碍的说话者的系统

来实现的，该系统包括：第一计算模块，适于自动确定说话者的输入话语的语音质量量度；第二计算模块，适于自动将语音质量量度与说话者的至少一个先前语音质量量度相比较；以及反馈模块，适于自动根据该比较来向说话者提供生物反馈。

本发明的目标还可以通过一种训练构音障碍的说话者的方法来实现，其中该方法包括以下步骤：自动确定说话者的输入话语的语音质量量度，自动将语音质量量度与说话者的至少一个先前语音质量量度相比较，以及根据该比较来自动向说话者提供生物反馈。

本发明的目标还可以通过一种在计算机中执行的计算机程序来实现，当在计算机中执行该计算机程序时，所述程序包括：用于自动确定说话者的输入话语的语音质量量度的计算机指令，用于自动将该语音质量量度与说话者的至少一个先前语音质量量度相比较的计算机指令，以及用于根据该比较来自动向说话者提供生物反馈的计算机指令。由此，根据本发明所需要的技术效果可以基于根据本发明的计算机程序指令来实现。

这种计算机程序可以保存在 CD-ROM 之类的载体上，或者可以通过因特网或另一种计算机网络得到。例如，在执行计算机程序之前，通过借助 CD-ROM 播放器从载体读取或是从因特网读取计算机程序，以及通过将计算机程序保存在计算机存储器中，可以将计算机程序加载到计算机中。特别地，计算机包括中央处理器（CPU）、总线系统、RAM 或 ROM 等存储器装置、软盘或硬盘单元等存储装置、以及输入/输出单元。可替换地，本发明的方法可以在硬件中实施，例如使用一个或多个集成电路来实施。

本发明的核心思想是提供一种借助简单廉价的训练系统来自动训练构音障碍的说话者的技术。例如，该思想可以借助标准的计算机系统、与计算机系统耦合的麦克风以及在计算机系统上执行的专用计算机程序来实现。

本发明的另一个核心思想是让说话者、例如正在康复的病人通过生物反馈效果来改善其状况。生物反馈是一种使用来自个人自身身体的信号对其进行训练，以便改善其身体状况的技术。在本例中，关于说话者的语音能力的信息被返还给说话者。优选地，该反馈信息取决于语音质量的改善或恶化。通过使用这个反馈信息，说话者可以逐步改善其语音能力。这种生物反馈机制允许说话者接收，例如查看或收听关于其语音能力的信息。为此目的，说话者的话语被“翻译”成了可以被说话者察觉的信号，例如闪光灯灯泡，蜂鸣器等等。例如，与先前结果相对比的语音质量的改善是

以“交通灯”或特定声音的形式来说明的。说话者尝试改善其技能并且监视其自己的成绩。根据通过生物反馈信号提供的结果，说话者将会“自动”调整其语音。通过使用这种生物反馈机制，详细的语言校正将变得不再必要。这样做极大降低了训练的成本和劳动。由于该训练完全是自动提供的，因此不会发生治疗或训练错误。由此，根据本发明的训练技术是非常可靠的。

根据本发明，说话者并不是被训练成产生只能为机器理解的语音。该目标很容易就可以通过请求说话者始终以相同方式重复某个单词来实现。取而代之的是，根据本发明的系统和方法旨在提高对于每一个平均听者、用户语音的智能性。因此，在本发明的优选实施例中，语音质量量度与所产生语音和平均正常语音生成的偏差是相互关联的。

本发明给出了用于训练构音障碍的说话者的语音质量的技术的一种实现方式。它使用户能够在没有语音治疗学家的情况下训练发音。而反馈将被提供给病人，并且该反馈将会指导其获得更好的语音质量。

通过本发明可以提供一种无人监管的训练。由此，本发明可以缩短乃至消除包括与语音治疗学家进行的昂贵的会话在内的成本密集的治疗。总而言之，本发明提供了一种在不需要监管的情况下训练构音障碍的说话者的简单、廉价、可靠的技术。

本发明的这些和其他方面将会根据从属权利要求中定义的后续实施例而被进一步阐述。

根据本发明的优选实施例，该系统包括适于记录说话者输入话语的记录设备。根据本发明的方法，说话者的输入话语是在确定语音质量量度之前记录的。该记录设备例如可以是麦克风，并且优选被连接到第一计算模块，以便直接在说话者的话音中馈送。可替换地，说话者的话语将会从麦克风传送到预处理设备，例如话音过滤器等等，其中所述预处理设备适于在借助第一计算模块执行的计算步骤之前预先处理说话者的话音。优选地，说话者的话语是在产生话语之后立即处理和评估的，换言之，该话语并未保存在数据库或类似设备中。由此，该训练系统可以用一种非常简单和廉价的方式来设计。

根据本发明的另一个优选实施例，该系统包括适于向说话者提供说话指令(speaking instruction)的接口模块。根据本发明的方法，说话指令是在记录说话者的输入话语之前提供给说话者的。为此目的，训练系统的接口模块优选包含一个用于显示将要由说话者重复的单词的监视器装置，

例如内置显示器。换句话说，说话者读取所要重复的单词。代替内置监视器装置，该接口模块也可以被连接到一个外部显示器。作为所述监视器装置的替换或补充，接口模块包括用于重放将要由说话者重复的单词的语音输出设备，例如内置扬声器。换句话说，说话者收听将要重复的一个或多个单词。与说话者随机说出某个单词的情形相比，如果实施这种说话指令，那么由第一计算模块执行的语音分析可以更加容易、快速，并且复杂度更低。由此，第一计算模块可以用一种复杂度更低的方式来内置，而这将会显著降低成本。

在将说话者的话语被输入训练系统之后，该系统将会自动评定语音质量。为此目的，在这里将会使用语音识别技术。

根据本发明的另一个优选实施例，第一计算模块包括适于在音位和/或单词级别上执行语音识别和语音分析的语音分析组件。根据本发明的方法，确定语音质量量度的步骤包括：在音位和/或单词级别上执行语音分析。由于构音障碍的说话者在大多数情况下在产生特定声音时有缺陷，因此，音位和/或单词级别上的语音分析是优选的。这些声音来自音位，也就是构成音节、单词和整个句子的声音差别的最小语音单位。单词级别的语音分析允许对相似单词中的音位质量进行比较：例如‘cheese’和‘freeze’都包含了‘e:’音位。此外还可以在相邻音位的上下文中分析音位：例如，如果病人在发‘killed’音的过程中显现出问题，但同时能够说出‘cold’、‘mold’和‘kitten’，那么明显是因为‘i’和‘l’的次序导致发生的问题。由此，以音位为基础的语音分析允许分析哪一个特定的声音偏离了多个非构音障碍的说话者发出的相同声音。

根据本发明的另一个实施例，第一计算模块包括适于将输入话语与参考数据相比较的比较组件，其中该比较是根据多个非构音障碍的说话者的数据来进行的。根据本发明的方法，确定语音质量量度的步骤包括：将输入话语与这种参考数据相比较。优选地，该参考数据是以多个非构音障碍的说话者的平均数据为基础的。为此目的，将会为在训练中使用的每一个单词提供参考发音数据，例如在数据库中提供。该参考发音数据在训练之前已经收集了，并且代表多个非构音障碍的说话者的平均单词发音。

根据语音分析以及与参考数据的比较，第一计算模块计算客观语音质量量度。该语音质量量度优选是为每一个单个训练单词确定的。所有结果，也就是训练单词的语音质量量度都会由第一计算模块保存在训练数据库中。优选地，该训练结果是连同时间戳一起保存的。借助这种手段，随后

可以将每一个结果分配给特定的日期和日时。

根据本发明的另一个优选实施例，第二计算模块包括适于相对于说话者的先前结果来执行语音质量量度的对比评估的评估组件。根据本发明的方法，比较语音质量量度的步骤包括：相对于说话者的先前结果来执行语音质量量度的对比评估。为此目的，使用历史训练数据，也就是相应单词的在前语音质量量度，来执行比较。这种历史训练数据保存在可以被第二计算模块访问的训练数据库中。优选地，第二计算模块不仅根据先前结果还根据执行训练的日时来评估语音质量。由此，时间相关的性能波动可以被最小化或者消除。

作为这种相对于先前结果来评估语音质量量度的结果，提供关于发音改善或损伤的信息，其中该信息优选是为每一个单独的单词提供的。换句话说，个人改善将被用于生物反馈。

在下文将会将会参考下列实施例和附图来举例详述本发明的这些和其他方面，其中：

图 1 显示了根据本发明的系统的示意性框图，

图 2 显示了根据本发明的方法的简化流程图。

在所描述的实施例中，用于训练构音障碍的说话者的系统 1 是借助例如笔记本或 PC 之类的标准计算机系统 2、与计算机系统 2 相耦合的麦克风 3 以及在计算机系统 2 上执行的专用计算机程序来实现的。在计算机系统 2 上执行的计算机程序实现的是第一计算模块 4 和第二计算模块 5。此外，计算机系统 2 包括组合的接口和反馈模块 6。系统 1 的所有装置都适于执行根据本发明的方法。所有模块、设备和组件都是以一种依照本发明的方法来运行数据和/或信息输入、处理和输出的过程的方式而被构造和编程的。特别地，计算机系统 2 的处理单元适于执行所有那些计算和运算输入数据以及确定和评定结果的任务。根据本发明，这种处理是借助专用计算机软件来实现的，其中该计算机软件包含了计算机指令，这些指令适于在计算机系统 1 的处理单元执行软件时执行本发明的方法。处理单元自身可以包括功能模块或单元，并且这些功能模块或单元是以硬件、软件或软硬件组合的形式实施的。

在第一步骤 101 中，用户（说话者）被要求说出某个单词或短句。为此目的，组合的接口和反馈模块 6 向说话者提供说话指令。该组合的接口

和反馈模块 6 包括计算机系统 2 的内置监视器 7 或者与计算机系统相连的外部监视装置，例如电视机。组合的接口和反馈模块 6 还包括用于重放将要由说话者重复的单词的内置扬声器 8。为了提供说话指令，所述组合的接口和反馈模块 6 是受专用计算机程序控制的。

在接下来的步骤 102，用户语音、也就是说话者的输入话语将会借助麦克风 3 而被记录，其中该麦克风 3 与计算机系统 2 相连，并且适于向计算机系统 2 传送语音信号，以便进行处理。

随后，说话者的输入话语的语音质量量度将会借助第一计算模块 4 而被计算。为此目的，在步骤 103 中通过第一计算模块 4 的语音分析组件 11 来执行语音识别。该语音识别在音位和/或单词级别上执行，以便确定所说的一个或多个单词。随后，第一计算模块 4 的比较组件 12 在步骤 104 中将输入话语与参考数据相比较，其中所述参考数据是以多个非构音障碍的说话者的平均数据为基础的。为此目的，比较组件 12 访问存储了每一个训练单词的参考数据的参考数据库 9。根据语音分析以及与参考数据的比较，第一计算模块 4 为每一个单独的训练单词计算客观的语音质量量度。

优选地，第一计算模块 4 是如下工作的：为发音训练给出单词 W。说话者的语音产生有可能是单词 O。现在，第一计算模块 4 计算两个概率。首先，给定语音产生 O 与给定单词 W 相匹配的概率，也就是所谓的“强制对准”。其次，给定语音产生 O 是任意音位序列的最大概率，也就是所谓的“自由音位循环”。这两个概率的比值（R=W/O）将会产生给定单词 W 的发音质量的量度。描述这种方法的详细资料可以在 H. Boulard 和 B. D'hoors 于 1994 年 4 月发表于 Proc. ICASSP 第 1 卷第 373 页的论文‘Optimizing Recognition and rejection performance in word spotting systems’中找到。

随后，语音质量量度 R 将被第一计算模块 4 连同时间戳以及相应的一个或多个单词一起保存在训练数据库 10 中。

在接下来的步骤 105 中，第二计算模块 5 的评估组件 13 自动将语音质量量度与说话者的至少一个先前语音质量量度相比较。在该步骤中，可以通过对当前会话的所有值求取平均，和/或通过使用该会话的最小/最大值，和/或通过在时间方面加权每一个值（最新近的值被加权最大）以及在此后求取平均，和/或通过对先前会话的累积结果加以考虑，由此顾及先前结果。为此目的，评估组件访问存储了每一个训练单词的先前结果的

训练数据库 10。如果该系统是首次使用的，那么测量得到的关于标准单词集合的语音质量将被用于对历史训练数据的数据库 10 进行初始化，并且还会用于对与所给出的生物反馈相关的性能等级进行初始化。

在接下来的步骤 106，根据当前练习结果来向说话者提供反馈。为此目的，组合的接口和反馈模块 6 提供了一个“信号灯”信令方案，在该方案中，如果有所改善，则显示绿灯，如果没有改善，则显示黄灯，如果出现恶化，则通过计算机系统 2 的内置监视器 7 来向说话者显示红灯。为了提供反馈信息，组合的接口和反馈模块 6 是由计算机程序来控制的。

优选地，说话者被要求重复训练处理，直至满足退出条件（步骤 107）。例如，这种退出条件可以是：达到最大训练时间，达到某一行的最大“绿”结果，达到某一行的最大“黄”结果，达到某一行的最大“红”结果，超出最大练习次数，手动停止。

一般的训练结果以及与每一个单独的单词或短语相关的训练结果被保存在结果数据库（未显示）中和/或被使得可以供治疗学家借助计算机系统 2 来使用，例如通过因特网之类的计算机网络针对治疗学家的在线数据传送来使用。

优选地，本发明既可以作为家庭中风测试的一部分来使用，也可以作为基于家庭的康复应用的一部分来使用，还可以在康复诊所中使用。此外，本发明还可以作为电子远程医疗平台的一部分来使用。

对本领域技术人员来说，很明显，本发明并不局限于上述说明性实施例的细节，在不脱离本发明的本质或是其基本属性的情况下，本发明是可以在其他特定形式中实施的。由此，本发明的所有方面都被视为说明性而不是限制性的，本发明的范围是由所附权利要求而不是前述说明书指示的，由此，落入权利要求含义和等价范围以内的所有变化全都落入其中。此外，非常明显的是，单词“包括”并未排除其他元件或步骤，单词“一”或“一个”并未排除多个，并且诸如计算机系统或另一个单元之类的单个元件可以满足权利要求中述及的若干个装置的功能。权利要求中的任何参考符号都不应该被解释成是对所涉及权利要求进行限制。

参考数字

1	系统
2	计算机
3	麦克风
4	第一计算单模块
5	第二计算模块
6	组合的接口和反馈模块
7	监视器
8	扬声器
9	参考数据库
10	训练数据库
11	语音分析组件
12	比较组件
13	评估组件

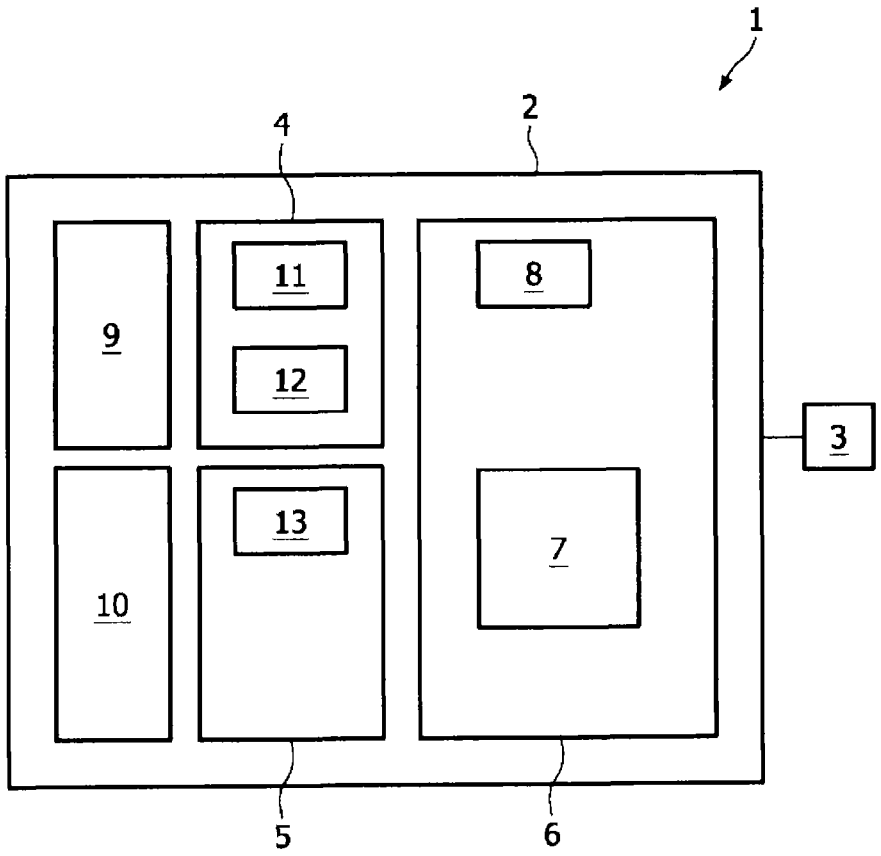


图 1

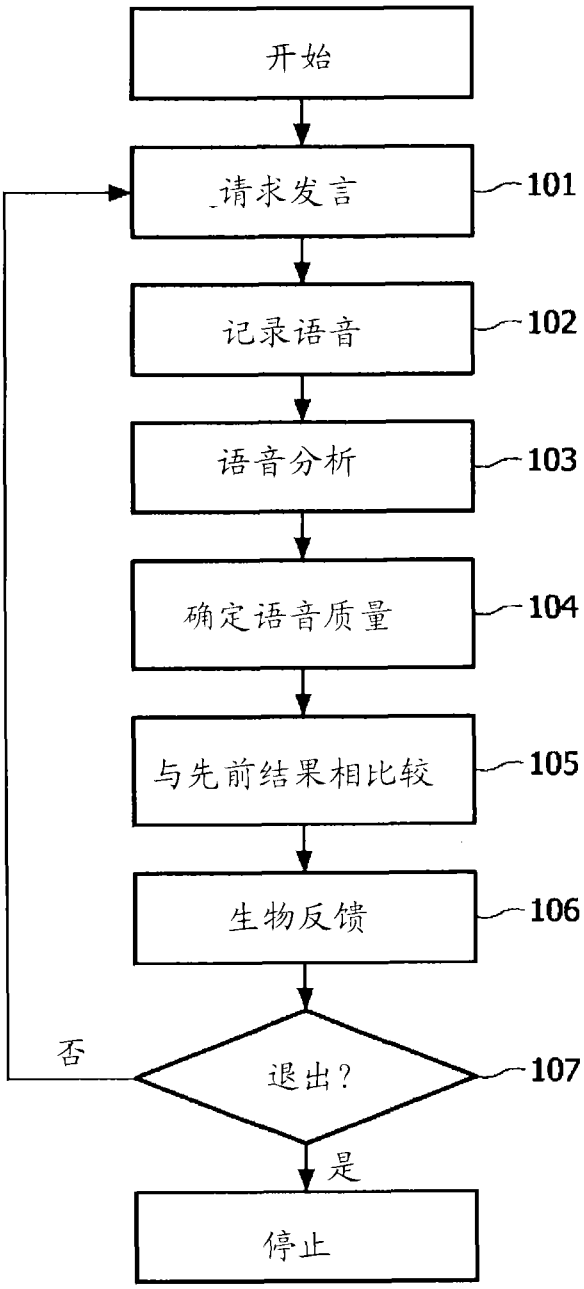

开始
请求发言
101
记录语音
102
语音分析
103
确定语音质量
104
与先前结果相比较
105
生物反馈
106
退出？
107
否
是
停止

图 2