

يستخدم التطبيق أربع تقنيات للنطق: التعرف التلقائي على الكلام، بينما يجب أن تكون الطرق المستخدمة لعلاج هذه الاضطرابات ممتعة لجذب المستخدمين (بشكل رئيسي الأطفال) الذين يعانون من اضطرابات في النطق. حيث يُمثل تبادل المعلومات بين هذه المكونات عبر السهام المختلفة. يستخدم فيكاليزا التعرف التلقائي على الكلام (ASR) لتمكين المستخدمين من تدريب أصواتهم، حيث يعرض التطبيق بعض المعلومات البصرية للمستخدم أثناء تشغيل تسجيل صوتي أو كلام مركب. تحتاج كل لعبة إلى نظام التعرف على الكلام الآلي لتعمل بشكل صحيح. كل إجابة هي كلمة معزولة. يجب على المستخدم اختيار الإجابة الصحيحة ونطقها بشكل صحيح لإكمال اللعبة بنجاح. تعمل هذه اللعبة مع الجمل، ويجب على المستخدم نطق جميع الكلمات في التسلسل الصحيح لإكمال اللعبة. كما هو موضح في الشكل رقم 1، تتلقى الألعاب معلومات حول الكلمات والألغاز والجمل التي يرغب المستخدم في تدريبها، يُحدد المستخدم الكلمات التي يريد نطقها وعدد مرات نطق كل كلمة، كما يوضح الشكل رقم 1، 3.3. معلومات المستخدم بما في ذلك جميع الكلمات والألغاز والجمل التي قد يواجهها المستخدم، مما يتيح للمستخدم إضافة كلمات وألغاز وجمل جديدة لتدريب نفسه، فقد تم تنظيم معلومات المستخدم بحيث يكون حملها سهلاً كما لو كان مجرد نسخ المجلد الذي يحتوي على جميع معلومات المستخدم من كمبيوتر واحد إلى آخر. فيوكاليزا: نافذة التدريب يمكن تكرار هاتين الخطوتين عدة مرات، التي تظهر في الشكل رقم 7. يتم الوصول إلى هذه النافذة من خلال القائمة الرئيسية وتتيح للمستخدم إضافة أو إزالة الكلمات أو الألغاز أو الجمل، إشارات الكلام يتم اكتسابها بتردد عينة قدره 16 كيلو هرتز وعمق بت بقيمة 16 بت. والميزات المستخدمة لنظام التعرف على الكلام هي MFCC 37 (معاملات التأثير المتغيرة بتردد الميل)، ليس التوليف الصوتي الطريقة الوحيدة المتاحة في فيوكاليزا ولكنها الأسهل للاستخدام: بمجرد إضافة كلمة جديدة أو جملة أو لغز إلى التطبيق، والتي في معظم الحالات ستألف من عدد قليل من النطقات بسبب عاملين: لا يستطيع أخصائيو النطق قضاء وقت طويل في تسجيل نطق كل مستخدم، فإن الدقة في نظام التعرف على الكلام الذي يعتمد على المتكلم باستخدام أساليب التقدير MAP تميل إلى أن تكون متساوية مع الدقة في نظام التعرف على الكلام الذي يعتمد على ML عندما يكون عدد النطقات كبيراً، يتم تخصيص مقياس لكل كلمة معترف بها، ويتم قبول أو رفض كل كلمة فرضية اعتماداً على مقياسها المقابل للثقة. يقيس تقريباً تحسن النطق لدى المستخدم. سيزيد مقياس الثقة المستخدم لكل كلمة في كل نطق كما يتحسن المستخدم مهارات التواصل لديه.