

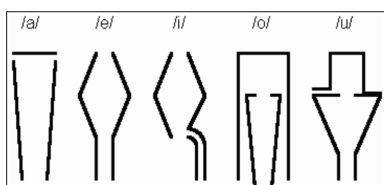
Synthétiseurs

Avant d'étudier plus en détail le synthétiseur de parole à partir du texte, reportons-nous quelque 200 ans plus tôt, à la genèse des machines parlantes, et marquons les jalons importants de cette histoire.

Les machines parlantes au fil du temps

L'histoire des machines parlantes a débuté il y a plus de 200 ans et plus exactement en 1779, date à laquelle Monsieur Christian Kratzenstein, professeur russe, expose les différences physiologiques qu'il a mises en évidence entre les 5 voyelles longues /a/ /e/ /i/ /o/ /u/. Encouragé par ses découvertes, il construit des appareils permettant de produire artificiellement ces voyelles. Ces appareils sont constitués de résonateurs acoustiques et de languettes vibrantes que l'on peut associer aux caisses de résonance et hanches ou cordes des instruments de musique.

La figure suivante représente les résonateurs de Kratzenstein ainsi que Mr Kratzenstein lui-même.



Résonateurs de Kratzenstein



Christian Gottlieb Kratzenstein (1723-1795)

Ceci correspond aux balbutiements des machines parlantes.

Depuis, deux siècles de recherches passionnées ont débouché sur des applications commerciales allant des plus ludiques aux plus utiles. L'histoire de la " machine parlante ", du 18ème siècle à nos jours, a été marquée par trois ères technologiques : l'ère de la machine parlante mécanique, l'ère de la machine parlante électrique et enfin, l'ère de l'ordinateur parlant.

Pour chacune de ces périodes, nous avons extrait les avancées les plus marquantes.

Machines parlantes mécaniques

Le premier qui fit réellement parler une machine fut Wolfgang Von Kempelen.

Von Kempelen avait pour domaine de prédilection l'étude de la production de la parole et était persuadé que cette connaissance approfondie lui permettrait d'en dégager des principes thérapeutiques. Il est considéré comme le premier phonéticien.

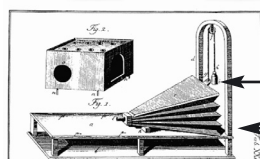
C'est en 1791 que paraît le livre dans lequel il décrit en détail, schémas à l'appui, sa " machine parlante ". Le but poursuivi par Von Kempelen était de donner tous les détails nécessaires afin qu'un tiers puisse, lui aussi, construire cette machine et l'optimiser.

Principe de base :

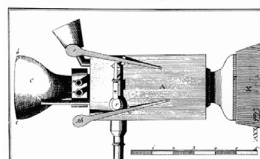
La "mécanique" de production de parole peut être décomposée en trois constituants principaux, à savoir :

- Les poumons et le diaphragme;
- Les cordes vocales;
- Le conduit vocal (cavités buccales et nasales).

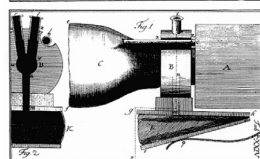
Le principe de la machine de Von Kempelen consiste à créer des éléments simulant ces différents organes afin de créer artificiellement de la parole. Les poumons furent simulés par un soufflet, le diaphragme par un contrepoids, les cordes vocales par les languettes vibrantes (sur le même principe que les résonateurs de Kratzenstein) et le conduit vocal par une boîte un peu sophistiquée. Mais voyez plutôt les illustrations tirées du livre de Von Kempelen !



Contrepoids (diaphragme)
Soufflet (poumons)



Boîte à vent (résonateur)



Extraits du livre de Von Kempelen

Synthétiseurs

II

Cette machine existe encore et est conservée au Deutsches Museum (von Naturwissenschaft und Technik) de Munich, dans le département des instruments de musique.



Machine de Von Kempelen (vue de plongée)



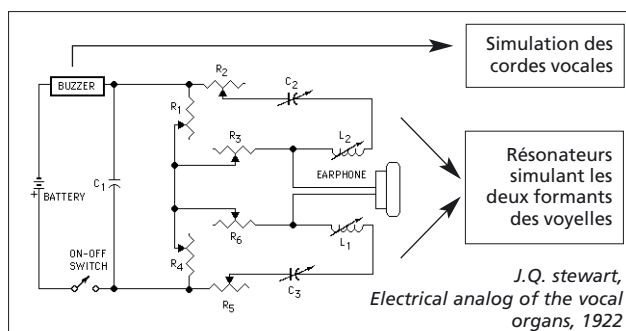
Machine de Von Kempelen (vue de profil)

Machines parlantes électriques

Le 20ème siècle fut marqué par les progrès notables en ingénierie électrique.

Ici commence la grande aventure de l'aspect "traitement du signal" de la production de parole artificielle, c'est-à-dire que c'est le signal de parole lui-même que l'on va tenter de reproduire plutôt que de reproduire l'appareil de production de la parole (conduit vocal, cordes vocales, ...).

En 1922, Stewart présentait le premier synthétiseur complètement électrique dont voici une représentation :



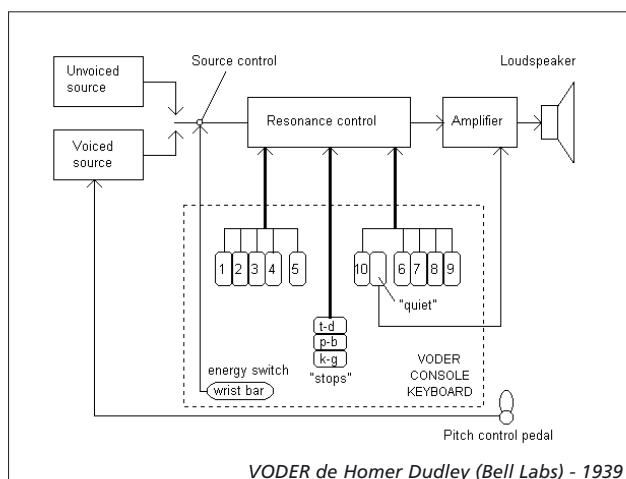
Ce synthétiseur permettait de synthétiser des voyelles en reproduisant deux formants¹ les caractérisant. Malheureusement, le résultat n'était pas très bon. En effet, Stewart modélisait deux formants alors que trois formants sont nécessaires à l'intelligibilité. Néanmoins, le ton était donné.

En 1939, alors qu'il étudiait le moyen de réduire la bande passante pour la transmission par téléphone dans le but de pouvoir faire cohabiter plusieurs communications sur une même ligne, Homer Dudley (Bell Labs) conçut le célèbre VOCODER (Voice Coder). Cette machine permettait de décomposer puis de recomposer un signal de parole. Ce fut le premier système de **codage de la parole**.

La même année, il présenta le VODER, machine capable de synthétiser de la parole.

L'utilisation de cette machine était néanmoins fort complexe et demandait un bon entraînement pour pouvoir l'utiliser.

Il est clair que les progrès de l'ingénierie électrique sont étroitement liés aux avancées basées sur l'aspect "signal de parole produit", par opposition à l'aspect "production de parole".



¹ Lors de la phonation (production du son humain), certaines fréquences du son produit par les cordes vocales sont amplifiées par les cavités de résonance (conduit vocal) en fonction de la fréquence de résonance propre à chaque cavité (cavités buccale, nasale, ...). Les fréquences qui sont amplifiées sont celles qui sont voisines de certains multiples de la fréquence caractéristique de la cavité, les autres fréquences étant affaiblies. Ce sont les fréquences renforcées que l'on nomme "formants". Les formants permettent de discriminer des voyelles ayant la même fréquence fondamentale, la même amplitude et la même durée. Les formants (abréviation F1, F2, F3, ...) permettent en fait d'identifier le timbre des sons.

Synthétiseurs



Dans les années 1950, Gunnar Fant (Royal Institute of Technology, Stockholm) met au point son synthétiseur pour voyelles baptisé OVE. Cette machine et ses développements furent à la base du système commercialisé INFOVOX.

Dans le même temps, en 1953, Walter Lawrence met au point le PAT (Parametric Artificial Taker)

qui permet une production des consonnes assez naturelle.

Dans les années 1960, la puissance de calcul augmentant, le traitement numérique du signal fit son apparition.

Ordinateurs parlants

Dans les années 1970, la technologie informatique a beaucoup évolué et a permis la synthèse de parole par ordinateur. A partir de ce moment, la parole de synthèse a pu être utilisée pour des applications bien précises, telle la lecture de textes par ordinateur.

Avec l'avènement de l'informatique, qui va de pair avec la création et la manipulation aisées de bases de données, est apparue l'idée d'accoler des sons courts pré-enregistrés plutôt que de les

générer artificiellement. C'est ce qu'on appelle la "synthèse par concaténation".

Ces techniques de synthèse font encore l'objet de bon nombre de recherches afin d'en améliorer le naturel de la parole sans perdre l'intelligibilité. De nombreux laboratoires de recherches y travaillent mais également des entreprises qui développent et commercialisent des produits tels que la lecture d'e-mails par téléphone, les centraux téléphoniques intelligents, les annonceurs automatiques et cetera.

