



Tout s'explique : **La synthèse vocale**

© Nic West

Peut-être avez-vous déjà croisé un utilisateur de synthèse vocale ? Si c'est le cas, vous vous êtes probablement demandé : « Mais qu'est-ce donc cette voix robotisée, à haut débit de parole et incompréhensible pour moi, que j'entends sortir du casque de mon voisin ? »...

Une synthèse vocale, qu'est-ce donc ?

La synthèse vocale, aussi appelée « text to speech » (TTS), est une voix artificielle obtenue grâce à un logiciel qui transforme un texte écrit en sons audibles.

A quoi sert une synthèse vocale ?

La synthèse vocale permet de rendre accessible, en toute autonomie, l'accès à l'information et à la communication à un grand nombre de personnes à capacités réduites.

A qui sert-elle ?

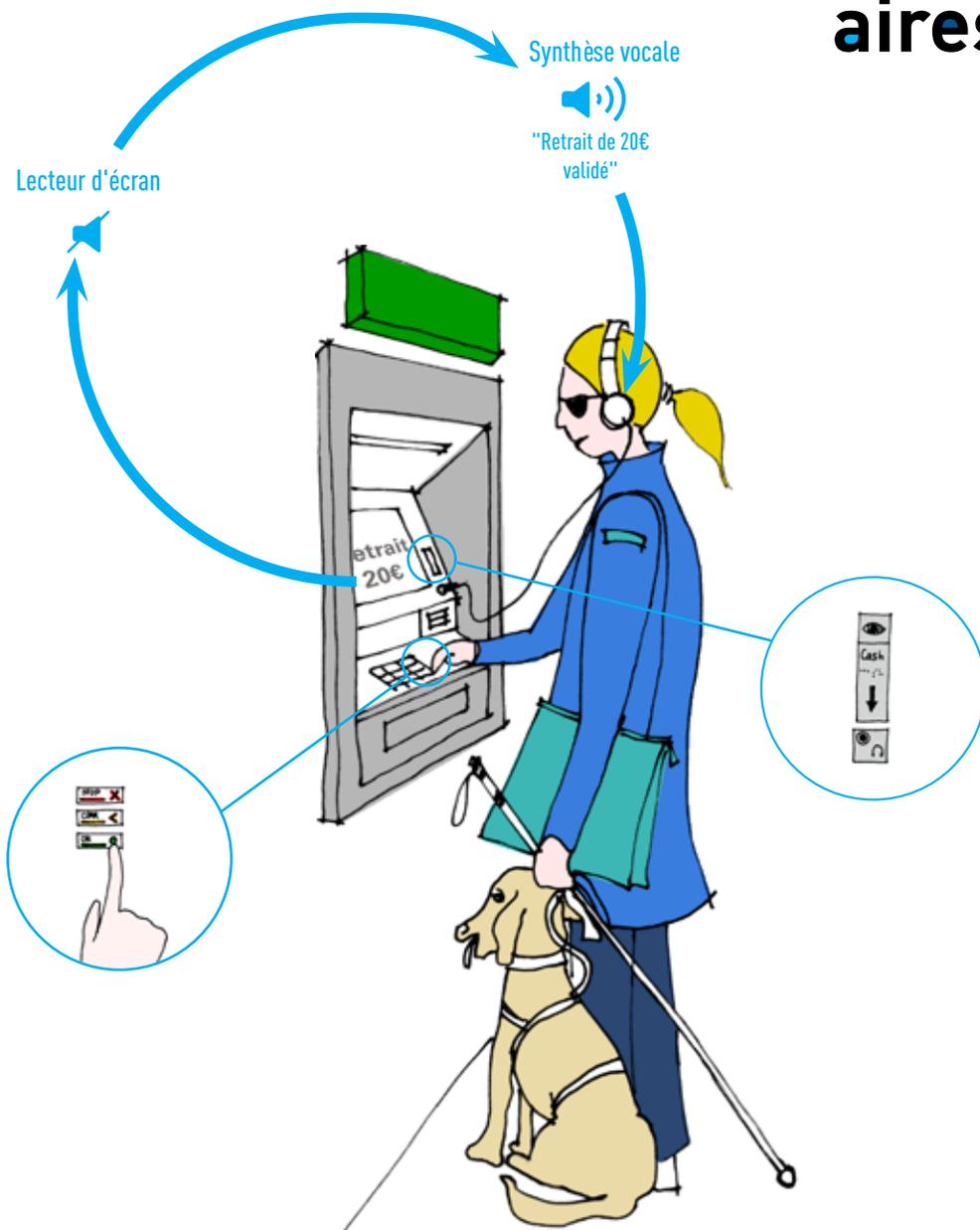
La synthèse vocale est principalement utilisée par ceux qui ne veulent ou ne peuvent utiliser leurs yeux (personnes déficientes visuelles).

Elle s'avère également indispensable à d'autres usagers :

- > les personnes dont la lecture à l'écran est trop difficile ou trop lente ;
- > les personnes âgées qui peuvent éprouver des difficultés de concentration ou de fatigue oculaire ;
- > les personnes avec des capacités de communication réduites ;
- > les personnes dyslexiques ;
- > les personnes peu ou pas scolarisées.

Ne pas confondre « synthèse vocale » et « reconnaissance vocale » !

La synthèse vocale est l'expression orale, par la machine, d'un écrit numérique alors que la reconnaissance vocale est la transformation, par la machine, d'un son émis par l'homme en un écrit numérique.



Concrètement, comment ça marche ?

Dans la mesure où la synthèse vocale lit à haute voix un message affiché sur un écran, il faut un **lecteur d'écran**. C'est un logiciel qui repère tous les éléments affichés et les hiérarchise en intégrant des informations sur le contenu (lien, niveau de titre, photo, champ à remplir, etc.).

Ce logiciel ne produit aucun son.

Le **synthétiseur vocal** reçoit les données traitées par le lecteur d'écran et les oralise.

Comment rendre cette voix artificielle la plus humaine possible ?

La difficulté pour les logiciels de synthèse vocale est de reproduire une rythmique naturelle. Pour y parvenir, trois facteurs doivent être pris en compte :

- > L'intonation apporte la touche mélodieuse selon le type de phrase (impérative, interrogative, affirmative).
- > Le rythme apporte la touche dynamique et nuancée (durée naturelle des syllabes).
- > L'intensité apporte la touche vivante (accents toniques, volume sonore).

Quel est le champ d'application de cette technologie ?

De nos jours, l'informatique et les nouvelles technologies sont couramment utilisées dans notre quotidien. Par conséquent, le champ d'application est très vaste. Voici quelques exemples :

- > la lecture d'e-mails et de SMS pour les personnes déficientes visuelles (et de manière plus générale, tout ce qui s'affiche sur un ordinateur) ;
- > les annonces dans les gares et les trains ;
- > les GPS ;
- > certains distributeurs de banque ;
- > ...