déterminisme, nous pouvons émettre l'hypothèse de l'existence d'un attracteur. Le signal change radicalement d'un phonème à l'autre, ce qui garantit, en théorie, que deux phonèmes se situent sur des hypersurfaces qui ne se coupent pas.

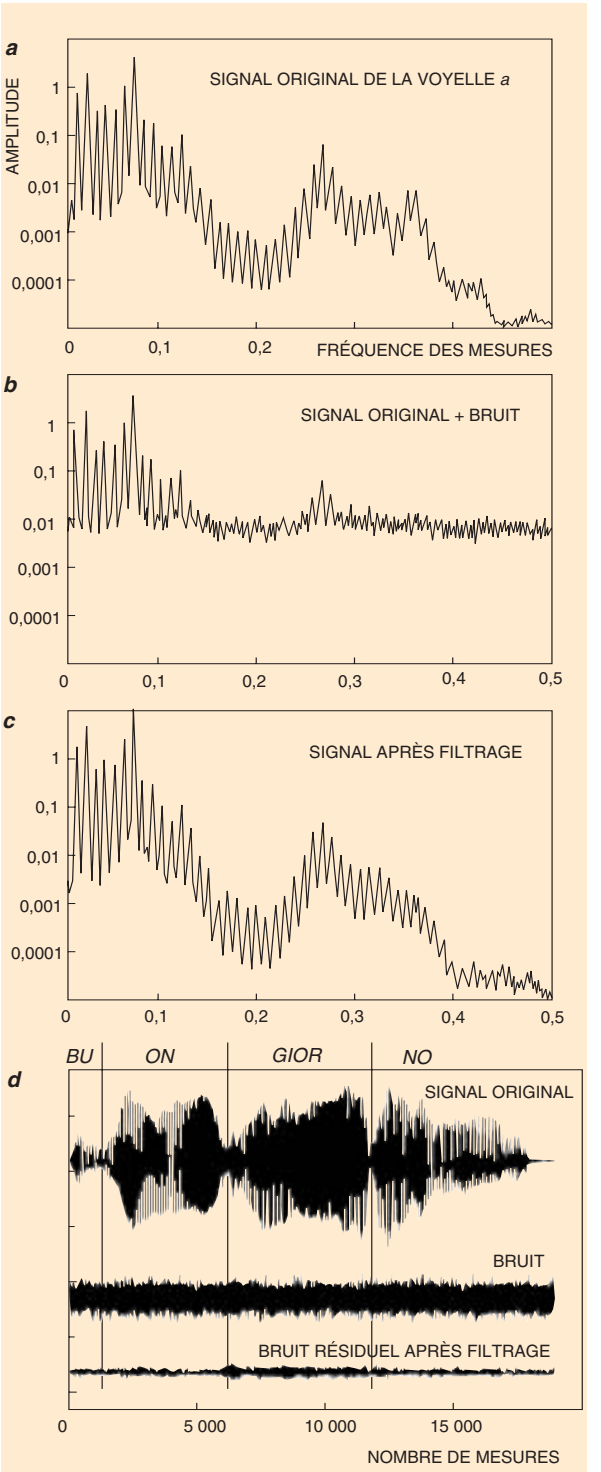
En pratique, les hypersurfaces caractéristiques des phonèmes ne se croisent pas, à condition que la dimension de l'espace des phases reconstruit soit suffisamment élevée. Quand on l'augmente, on «déplie» les hypersurfaces et elles ne se recoupent pas avec elles-mêmes. Ici, on augmente cette dimension de façon à éviter que deux phonèmes différents soient représentés par une même structure dans l'espace des phases reconstruit. Cette condition est importante pour retrouver le signal vocal. En effet, quand la trajectoire présente des intersections, le système ne peut être déterministe, car, si le point d'intersection constituait par exemple une condition initiale, le système pourrait évoluer aussi bien dans une direction que dans une autre, ce qui est contraire à l'idée de déterminisme.

Simple comme bonjour

Pour reconstruire les attracteurs caractéristiques de chaque phonème, nous avons observé que chaque vecteur devait couvrir une fenêtre temporelle de longueur équivalente au moins à un sous-phonème. En pratique, on enregistre un point de mesure s_k, environ toutes les 0,045 milliseconde (échantillonnage à 22,05 kilohertz). La longueur d'un sous-phonème varie entre 80 et 150 points, soit environ 3,5 et 6,7 millisecondes. Il suffit de choisir le délai temporel τ de telle sorte que $m\tau$ soit de l'ordre de grandeur de la longueur d'un phonème. D'après de récentes études empiriques, la dimension m associée à la phonation prolongée d'une voyelle est de trois pour une voix normale et de cinq pour un sujet affecté d'une pathologie des cordes vocales. En revanche, une phrase entière est trop complexe pour être analysée facilement. Nous supposons cependant que la dimension impliquée dans un sous-phonème ne dépasse pas dix. D'après le théorème de F. Takens, le choix d'une dimension m égale à deux fois la dimension obtenue pour un phonème augmentée de quelques unités, devrait permettre de disposer d'un espace des phases reconstruit dans lequel les phonèmes pourront être distingués sans ambiguïté.

En analysant plusieurs phrases, et en comparant les résultats du filtre réalisé avec différents délais temporels et différentes dimensions, nous avons établi que l'espace des phases reconstruit optimal est obtenu en prenant un délai temporel égal à l'intervalle séparant cinq points consécutifs de la série temporelle, soit 0,18 milliseconde. Nous avons choisi une dimension m égale à 25. Tous nos résultats ont été obtenus avec ces deux valeurs pour les paramètres τ et m.

Pour vérifier les performances du filtre, nous avons effectué une série d'expériences, de plus en plus complexes. Nous avons commencé par l'étude de la phonation prolongée de la voyelle *a*. Nous avons d'abord enregistré un signal en l'absence de bruit *(voir la figure 4)*, puis nous l'avons perturbé au point de détruire beaucoup des structures initialement présentes. Nous avons ainsi obtenu un signal aplati, caractéristique d'un bruit «blanc», c'est-à-dire pour lequel la densité spectrale de puissance est constante quelle que soit la fréquence. En appliquant le filtre, nous avons identifié presque toutes les structures de départ, bien qu'elles se trouvent au-dessous du niveau du bruit (leur amplitude est inférieure à celle du bruit). Peu importe à ce stade de comprendre



4. LE FILTRE est d'abord validé sur une voyelle soutenue : le signal d'un *a* émis pendant trois secondes est enregistré : il forme une série temporelle *(a)*. Un bruit numérique est ajouté et perturbe le signal *(b)*. Le signal est restauré après le filtrage *(c)* : la quasi-totalité des structures présentes dans le signal original réapparaissent. Puis le filtre est mis à l'épreuve avec le signal *bonjour (buon giorno) (d)*, auquel est ajouté un signal de bruit *(au centre)* ; on obtient ainsi le signal d'entrée du filtre. Le signal obtenu à la sortie du filtre est comparé au signal vocal original : la différence est un bruit faible, considéré comme un bruit résiduel *(en bas)*.