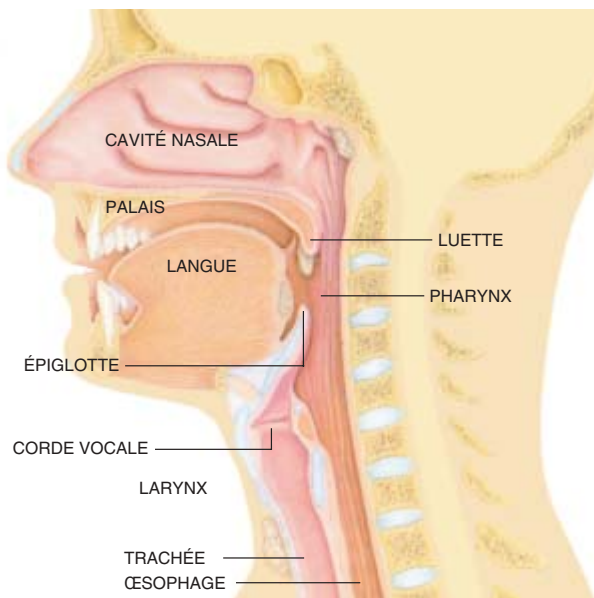




2. LES ORGANES DE LA PRODUCTION VOCALE sont nombreux. La voix est produite par l'air en provenance des poumons qui traverse des résonateurs (le pharynx ; la cavité buccale ; la cavité labiale ; les fosses nasales). L'air entre en vibration au contact des articulateurs rencontrés sur son parcours : d'abord le larynx, armature de cartilage qui abrite les cordes vocales (des ligaments recouverts d'une muqueuse qui fait vibrer l'air) ; puis l'air rencontre le pharynx, le voile du palais, la langue, les lèvres... Les positions des articulateurs modifient la forme du conduit vocal et le timbre du son produit.

une erreur de prévision croît aussi rapidement : la sensibilité aux conditions initiales réduit drastiquement la durée sur laquelle les prévisions sont fiables. Finalement, la détermination du comportement à long terme d'un système chaotique est aussi complexe que celui d'un système aléatoire.

De nombreux outils mathématiques ont été développés pour traiter des phénomènes chaotiques, ou non linéaires en général, et ils trouvent de nombreuses applications, dans le traitement du signal de la voix par exemple. Lorsqu'un signal vocal est enregistré, il est inévitablement pollué par du bruit, qui provient de l'environnement sonore ou d'une mauvaise réception. Ce signal parasite se mêle au signal vocal et le perturbe, jusqu'à le rendre méconnaissable parfois. Les physiciens développent des filtres pour éliminer les bruits parasites et améliorer les performances des dispositifs de reconnaissance vocale ou celles des amplificateurs, par exemple : la réduction du bruit dans le traitement de la parole présente un intérêt considérable dans les télécommunications et dans les interfaces vocales homme/machine.

Les filtres utilisent des techniques de traitement du signal très variées, plus ou moins efficaces pour faire la distinction entre le bruit et le signal lié à la parole dans les signaux enregistrés. Nous avons mis au point un filtre qui exploite des outils issus de la théorie du chaos et qui repère, puis sépare le bruit du signal vocal. L'astuce consiste à différencier le bruit – composé de signaux d'origine aléatoire – des signaux vocaux qui, bien que complexes, ont des caractéristiques déterministes. En pratique, le signal sonore est enregistré par un microphone, puis transformé en une série de grandeurs scalaires, lesquelles sont transformées en vecteurs, puis traitées à l'aide des techniques de la théorie du chaos, afin d'en extraire le signal exempt de bruit parasite.

Les phénomènes responsables de la production vocale sont non linéaires. La voix est produite par le passage de l'air dans le conduit vocal. Les organes impliqués sont nombreux : le larynx (où se trouvent les cordes vocales), les fosses nasales, le pharynx, l'épiglotte, la langue, le palais, les lèvres et même les dents (voir la figure 2). Quand l'air passe des poumons vers la bouche, la position des organes change afin de produire différents sons. Par exemple, quand l'air circule librement dans la glotte, l'espace entre les cordes vocales, on prononce une voyelle ; en revanche, quand le passage de l'air est obstrué, entièrement ou en partie, on émet une consonne. La mise au point d'un modèle de la production vocale humaine est aujourd'hui irréaliste.

Une alternative consiste à analyser les propriétés du système vocal : la parole émise est une onde, dont le signal est représenté par son amplitude en fonction du temps. Le signal d'une phrase ou même d'un mot, enregistré par un micro, est complexe et variable (voir la figure 1). Il varie peu par endroits, et beaucoup à d'autres, où son amplitude est supérieure. Pourtant, en y regardant de plus près, on peut repérer des structures approximativement stationnaires (leurs propriétés sont invariables dans le temps) et moins complexes : ce sont les phonèmes (que l'on classe en consonnes, en voyelles et en semi-voyelles, par exemple *ou*, *on*...). Le nombre de phonèmes varie en fonction de la langue (des dialectes africains en comptent jusqu'à 60). À l'intérieur de chaque phonème, le signal est presque périodique (voir la figure 3). Le phonème est caractérisé par la répétition de sous-structures, dites sous-phonèmes. Toutefois, deux phonèmes consécutifs sont très différents, si bien qu'on ne peut prévoir l'état futur du signal.

La trajectoire dans une hypersurface

Dans un système purement déterministe, quand l'état présent du système est connu, les états futurs sont déterminés. Dans ce cas, l'identification des grandeurs physiques qui décrivent le système, c'est-à-dire des variables indépendantes, est cruciale. Considérées dans leur ensemble, elles prennent le nom de « vecteurs d'état » du système ; chaque vecteur d'état – ce qui revient à dire chaque état du système – est représenté de façon univoque par un point dans un espace dit vectoriel, aussi nommé espace des phases ou espace d'états. Il existe une équivalence entre l'état du système et un point de cet espace vectoriel. Cet espace mathématique comporte autant de dimensions qu'il y a de variables indépendantes dans le système dynamique (ses axes de coordonnées correspondent à ces variables indépendantes). Par exemple, il en comporte six pour un ballon (trois coordonnées spatiales et trois coordonnées de vitesse). L'espace des phases sert à représenter l'évolution des systèmes dynamiques : on détermine la dynamique du système réel en étudiant la structure de la trajectoire représentative de l'évolution du système dans l'espace des phases (voir l'encadré de la page 82).

En général, un système déterministe dynamique est défini par un ensemble d'équations différentielles, pour lesquelles une solution unique existe à partir de conditions initiales données. Cela se traduit par l'existence et l'unicité d'une trajectoire, c'est-à-dire une succession de points (ou encore de vecteurs d'état) dans l'espace des phases correspondant, où chacun est déterminé de façon univoque par le précédent. Le déterminisme inhérent au système contraint cette trajectoire à se trouver dans une

hypersurface de l'espace des phases (une hypersurface dans un tel espace multidimensionnel correspond à une surface dans un espace tridimensionnel). Dans cette situation, la trajectoire prend le nom d'attracteur.

D'ordinaire, au cours d'une expérience, on n'observe ni l'espace des phases ni un système d'équations différentielles. Le résultat des mesures est une série temporelle, c'est-à-dire une succession de nombres, qui doit être convertie en vecteurs d'état. Cette étape, dite de reconstruction de l'espace des phases, se résout grâce à la méthode de plongement. On ne construit pas ainsi l'espace original des phases, mais un espace équivalent, l'espace des phases reconstruit. Les deux espaces peuvent être mis en correspondance biunivoque (un élément de l'ensemble d'arrivée correspond à un seul élément de l'ensemble de départ, et réciproquement).

Considérons une série temporelle obtenue d'après l'enregistrement de paroles à l'aide d'un microphone. Les différences de pression, provoquées par la voix, sont transformées en un signal électrique dont on enregistre l'intensité I régulièrement : les points de mesure sont séparés par un petit intervalle de temps dt constant. Supposons que l'ensemble de la série temporelle soit constitué de N points de mesure, c'est-à-dire N valeurs de l'intensité du courant délivrée par le microphone. La série temporelle est donc représentée par la suite $\{I(t)\}$ où t varie entre t_0 , l'instant de la première mesure, et $t_0 + Ndt$. Par ailleurs, pour définir complètement un état du système de production de la parole humaine, il faut m variables indépendantes.

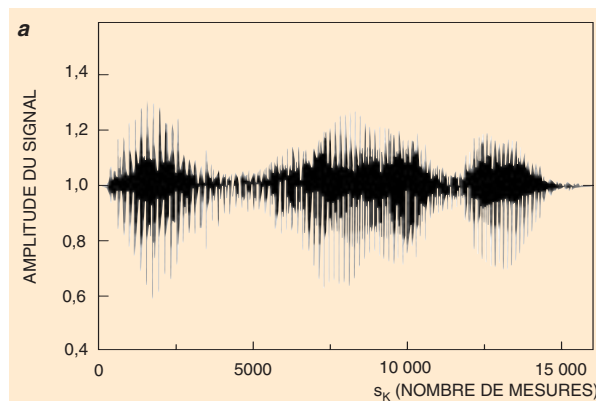
Ces m variables peuvent être obtenues à partir de la seule mesure de l'intensité du courant $I(t)$ grâce à une technique dite des coordonnées décalées temporellement (ou méthode des retards), c'est-à-dire de la valeur $I(t)$ à l'instant t , de la valeur $I(t + \tau)$ à l'instant $t + \tau$, et ainsi de suite jusqu'à disposer de m coordonnées qui définissent l'état du système (soit un point) dans l'espace reconstruit de dimension m . Ici τ correspond à un décalage temporel, multiple de l'intervalle dt entre deux points de mesure consécutifs. Cette technique a été proposée par l'Américain Norman Packard et ses collègues (sur une idée du mathématicien David Ruelle) en 1980. En répétant l'opération pour chaque instant t , on obtient une succession de points dans l'espace reconstruit, qui définit ainsi une trajectoire représentative de l'évolution du système.

En 1981, Floris Takens, mathématicien hollandais à l'Université de Groningen, a établi un théorème qui stipule que, si m est supérieur à deux fois la dimension de l'objet étudié, alors l'espace des phases reconstruit possède toutes les propriétés de l'espace original des phases. Par exemple, une simple boucle est un objet de dimension 1 ; il faut au moins un espace de dimension 3 pour être certain de pouvoir « plonger » cette boucle dans un espace reconstruit à partir de la mesure de l'une de ses coordonnées. En outre, toutes les grandeurs dynamiques sont invariantes. La condition imposée par ce théorème est suffisante, mais pas nécessaire, dans le sens où, dans certains cas, une dimension m plus petite peut suffire. Dans le cas de la simple boucle, un espace reconstruit de dimension 2 est parfois suffisant. Notons que les attracteurs chaotiques ont une dimension non entière, nommée dimension fractale.

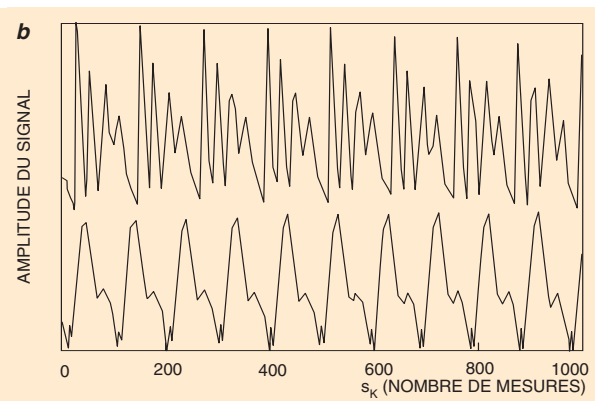
Déformation de l'attracteur

Considérons de nouveau un système déterministe, et laissons-le évoluer à partir d'un état initial. Dans l'espace des phases correspondant, nous obtenons, comme nous l'avons vu, une trajectoire (ou attracteur) qui se trouve sur une hypersurface. En revanche, un système stochastique, c'est-à-dire dans lequel se trouve une composante non déterministe, ne présente pas cette propriété. L'évolution d'un système stochastique étant aléatoire, les points, obtenus par la méthode des retards, se répartissent au hasard dans l'espace des phases reconstruit, quelle que soit sa dimension. En effet, un système complètement aléatoire ne peut être décrit par un nombre fini de variables indépendantes. Par conséquent, il est impossible d'imposer une limite supérieure à la dimension de l'espace des phases. Grâce à cette discrimination, on différencie un signal déterministe d'un signal stochastique.

En pratique, quand un signal vocal x_k est contaminé par un bruit n_k , il mène à la série temporelle $s_k = x_k + n_k$. Chaque point de l'espace des phases subit alors une déviation de sa position originale. La trajectoire n'est plus unique et elle n'appartient plus à une hypersurface. En d'autres termes, l'attracteur se déforme : l'introduction du bruit provoque la formation d'un nuage de points autour de l'attracteur initial.



3. L'AMPLITUDE DU SIGNAL VOCAL est enregistrée à raison d'un point toutes les 0,045 millisecondes (échantillonnage à 22,05 kilohertz). La série temporelle qui en résulte (a) (environ 15 000 points) dure 0,68 seconde. Quand on agrandit la représentation du signal vocal, des structures périodiques distinctes y apparaissent (b). Elles forment



différents phonèmes, composés de sous-structures (ou sous-phonèmes) qui se répètent. Le phonème inférieur présente dix sous-phonèmes, dont la durée est d'environ 4,5 millisecondes. Le phonème supérieur présente huit sous-phonèmes. Le filtre différencie le bruit (d'origine aléatoire) du signal vocal composé de ces structures quasi périodiques.