

Paradigme catastrophique et perception catégorielle*

JEAN PETITOT *École des Hautes Études en Sciences Sociales, Paris*

En phonétique, les phénomènes de perception catégorielle permettent de comprendre le rapport existant entre le flux audio-acoustique et le code phonologique, c'est-à-dire, pour reprendre les termes hjelmsleviens, entre l'organisation de la *substance* de l'expression phonétique et la structure relationnelle de sa *forme*. Leur compréhension théorique est donc d'une importance particulière pour les sciences phonétiques et, plus généralement, pour les sciences structurales puisque, d'après le principe de parallélisme entre la dimension de l'expression et celle du contenu, l'organisation paradigmatique des phonèmes sert de 'modèle' pour l'organisation paradigmatique en général. Dans cette perspective, nous nous proposons de montrer que ces phénomènes sont des cas perceptifs de phénomènes catastrophiques et qu'il est en conséquence souhaitable d'introduire les outils conceptuels et mathématiques de la théorie des catastrophes dans la méthodologie phonétique.

1 LA PERCEPTION CATÉGORIELLE

1.1 Définition.

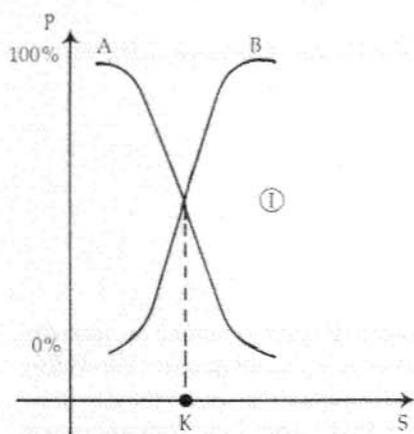
Découverte en 1957 par A. Liberman, la perception catégorielle s'oppose à la perception continue. Considérons un 'continuum' de stimuli allant d'une syllable $S_1 = C_1V$ à une syllable $S_2 = C_2V$ de même voyelle, les consonnes C_1 et C_2 (par exemple des occlusives) ne différant que par un seul indice acoustique (par exemple par le voisement comme dans [ba]/[pa], [du]/[tu] etc. ou par la place d'articulation comme dans [bo]/[go], [pi]/[ti] etc.). Ce 'continuum' est en fait une suite discrète de N stimuli (N étant de l'ordre d'une dizaine) dont les premiers et les derniers sont naturels (articulatoirement productibles) et les autres, intermédiaires, synthétiques. À partir de ce matériel, on soumet un groupe de sujets à des tests d'identification et de discrimination (par exemple par la méthode ABX).

* Ces réflexions ayant été inspirées par la lecture de *La Charpente phonique du langage*, nous aimerions les dédier à la mémoire de Roman Jakobson et à Linda Waugh. Cet article développe une intervention préparée pour le *Colloque Transdisciplinaire sur la Glossogénétique* qui a été tenu à Paris à l'UNESCO en 1981. La version italienne doit paraître dans PROMETEUS International.

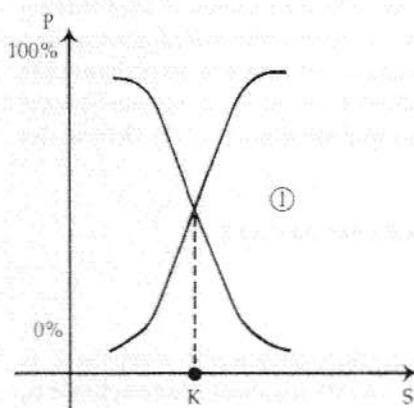
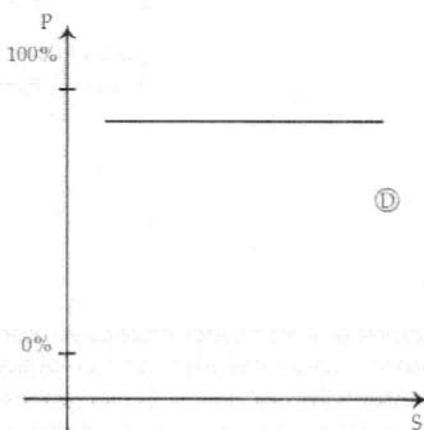
FIGURE 1 P = pourcentage, S = stimuli

(a) Perception continue (I = identification, D = discrimination, K = interface entre deux catégories).

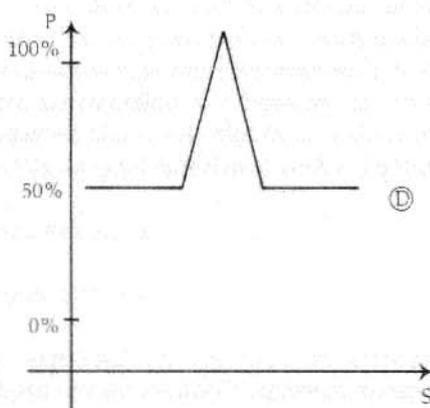
(b) Perception catégorielle.



(a)



(b)



L'on constate alors qu'il n'existe pas de discrimination intracatégoriale : les sujets ne discriminent deux stimuli voisins n et $n + 1$ que s'ils sont situés de part et d'autre de l'interface séparant deux catégories identifiables adjacentes. Autrement dit, et contrairement à ce qui se passe pour une perception continue comme celle des couleurs, la discrimination est subordonnée à l'identification et se fait sur des bases absolues, non relatives (cf. Fig. 1). Ainsi que l'ont noté M. Studdert-Kennedy et A. Liberman, 'categorical perception refers to a mode by which stimuli are responded to, and can only be responded to, in absolute terms' (1970 : 234).

1.2 *Fonction.*

L'importance fonctionnelle de la perception catégorielle est évidente. C'est en effet l'absence de discrimination intra-catégoriale qui permet de comprendre la *discrétisation perceptive* du flux audio-acoustique, permettant d'en faire le support du code phonologique. Cette discrétisation porte essentiellement sur les consonnes (et plus particulièrement sur les occlusives) c'est-à-dire sur les phonèmes encodés dans le flux (la perception des voyelles et des fricatives est par exemple plus continue que catégorielle). Les phonèmes encodés sont catégoriques en tant qu'immédiatement donnés à la perception et tout porte donc à croire qu'il existe un mode spécifique (a 'speech mode') de traitement et de décodage qui leur est associé (Liberman *et al.* 1967). La perception catégorielle est bien spécifique puisque si, étant donné un continuum syllabique $C_1V - C_2V$ perçu catégoriellement, on extrait la partie consonantique (transitions des formants et bruit d'explosion) on obtient un continuum de bruits non phonétiques ('chirps') dont la perception est, elle, continue.

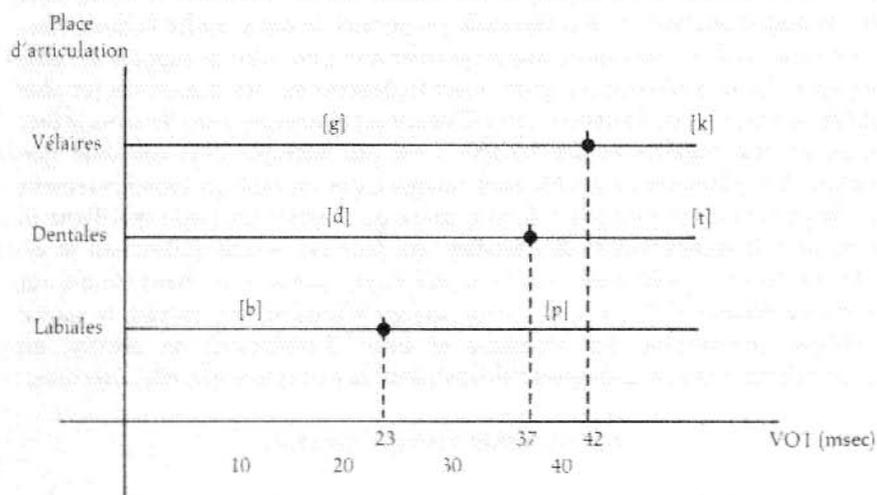
1.3 *Situation abstraite générale.*

D'une façon générale, les phénomènes phonétiques de perception catégorielle sont dus à la façon dont les indices acoustiques *contrôlent* les percepts. À ce titre, il relève de la situation abstraite générale suivante. Soient (u_1, \dots, u_r) des paramètres (en l'occurrence des indices acoustiques) parcourant un espace W et contrôlant les états internes d'une 'boîte noire' S (en l'occurrence un système perceptif). Ce qu'il s'agit de comprendre est la façon dont un tel système contrôlé peut *catégoriser son espace de contrôle*. Il s'agit là de situations très différentes de celles décrites par la théorie des automates. En effet, au lieu d'être en présence d'un ensemble discret d'inputs et d'un ensemble discret d'outputs engendrés, en fonction des inputs, par des transitions entre des états internes eux-mêmes discrets, on se trouve en présence d'un ensemble *continu* W d'inputs ayant valeur de contrôle, les transitions entre les états internes ayant pour effet non pas d'engendrer des outputs mais d'induire un système K d'interfaces, de seuils, de discontinuités dans l'espace externe W . Notons qu'il existe des cas *physiques* particulièrement typiques de cette situation générale, à savoir les phénomènes de transitions de phases. En ce sens, il est légitime de traiter (d'abord analogiquement puis ensuite théoriquement) la perception catégorielle comme l'induction de diagrammes de phases dans les espaces d'indices acoustiques contrôlant les percepts.

1.4 *Exemples.*

De nombreuses expériences ont été menées ces dix dernières années sur la perception catégorielle. On a en particulier intensivement étudié les interfaces K induites sur l'axe du VOT ('voice onset time' i.e. indice du voisement) par l'identification des paires fondamentales d'occlusives [b]/[p], [d]/[t] et [g]/[k]. Les expériences pion-

FIGURE 2 : Insuffisance des expériences de Lisker et Abramson : la donnée de 3 points ne suffit pas à reconstruire la morphologie des interfaces classifiant les occlusives relativement aux paramètres de voisement et de place d'articulation.

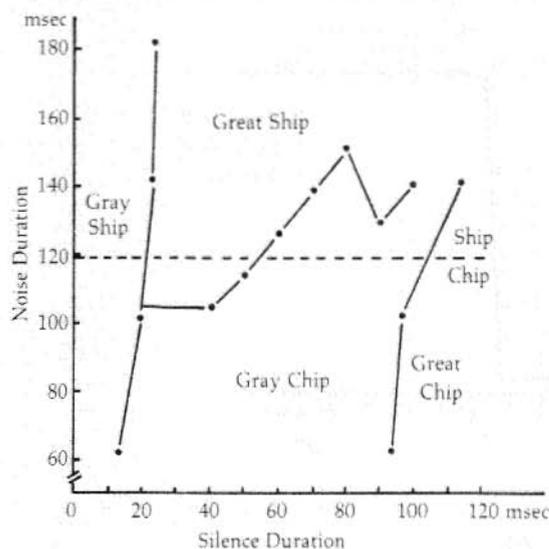


nières dans ce domaine ont été réalisées en 1970 par Lisker et Abramson qui ont analysé la variation de K en fonction du point d'articulation (cf. Fig. 2).

Mais ces expériences sont largement insuffisantes. En effet la place d'articulation dépendant (comme le voisement) d'indices acoustiques *continus* (comme par exemple la fréquence du bruit d'explosion ou la transition du second formant, cf. la théorie du locus de P. Delattre), le système d'interfaces K induit par la perception catégorielle catégorise un espace externe W multidimensionnel de dimension r . Or K devant *classifier* et *discriminer* les percepts contrôlés par W , il doit décomposer K en domaines (en catégories). Cela exige *a priori* qu'il soit de codimension 1 (i.e. de dimension $r - 1$). Qui plus est, l'information fondamentale est l'information géométrique fournie par sa morphologie. Or, comme il apparaît clairement dans la Fig. 2, les résultats de Lisker et Abramson ne permettent pas de reconstruire une morphologie de codimension 1 (i.e. de dimension $2 - 1 = 1$) dans l'espace externe des indices du VOT et de la place d'articulation.

Certains essais de reconstruction explicite d'un diagramme de phase dans un espace de contrôle acoustique ont cependant été menés à bien. Un exemple en est fourni par les expériences de B. Repp sur les fricatives en anglais (Repp *et al.* 1978). Repp considère deux paramètres de contrôle: une durée de silence ΔS et une durée de bruit de friction ΔB et analyse leur intégration dans la discrimination des fricatives et des affriquées. Dans le cas d'un énoncé comme 'did anybody see the gray ship', l'espace externe (ΔS , ΔB) se trouve catégorisé en 4 domaines correspondant aux perceptions [gray ship], [gray chip], [great ship] et [great chip] (cf. Fig. 3).

FIGURE 3 : D'après Repp *et al.* (1978). Boundaries that divide the several response categories, represented as joint functions of duration of silence and duration of fricative noise.



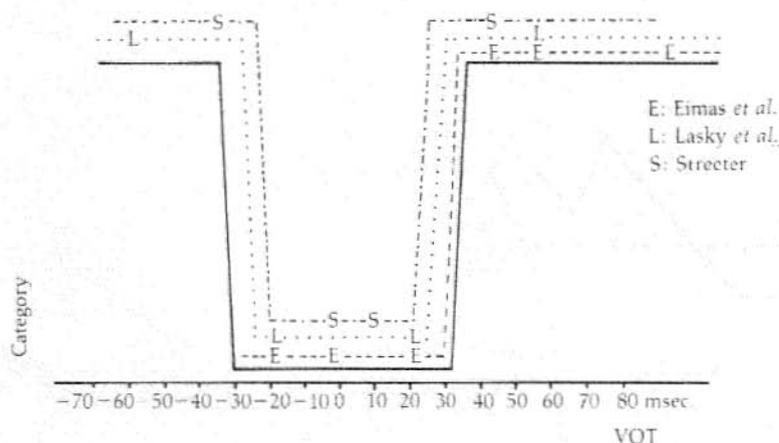
1.5 Spécificité.

Contrairement à ce que l'on a cru au début des recherches, la perception catégorielle n'est pas spécifiquement phonétique. Elle existe de façon assez générale dans le domaine sonore. Par exemple, la perception des timbres musicaux est catégorielle (Cutting et Rosner 1974). Il en va de même de la perception des intervalles musicaux chez les musiciens professionnels possédant 'l'oreille absolue' (Siegel et Siegel 1977). Comme l'on sait que les musiciens professionnels traitent l'information préférentiellement dans l'hémisphère gauche (dominance de l'oreille droite testable par des expériences d'écoute dichotique), cela laisse supposer qu'il existe dans cet hémisphère un mode catégoriel spécifique du traitement de l'information ayant pour fonction de la discrétiser (de la digitaliser) et de la transformer en code.

Il existe aussi des phénomènes de perception catégorielle ne concernant que l'organisation temporelle et concernant donc un niveau très abstrait. J. Mehler a par exemple montré que si l'on distribue trois battements 1, 2 et 3 sur un intervalle de 600 ms le battement 2 étant situé dans une position intermédiaire variable quelconque, la perception est catégorielle et partitionne les stimuli en trois classes correspondant respectivement aux perceptions invariantes $1-2/3$, $1/2/3$, et $1/2-3$, les interfaces se situant à environ $\pm 20-30$ ms de la position centrale où le battement 2 est à 300 ms (Mehler, Bertoini, 1980).

Il existe également des phénomènes de perception catégorielle dans le domaine visuel. Un exemple typique en est fourni, pour des stimuli lumineux intermittents, par l'existence d'un seuil sensoriel au-delà duquel les stimuli sont perçus comme continus (seuil de 'flicker-fusion').

FIGURE 4: La catégorisation universelle du VOT (Voice onset time) chez les bébés. D'après Mehler, Bertoncini (1980). Lasky (1975) concerne des bébés espagnols et Streeter (1976) des bébés Kikuyu.



1.6 Innéité.

Revenons à la catégorisation de l'axe du VOT. Si l'on expérimente sur des langues comme l'anglais ou le français où il n'existe que deux possibilités de voisement (voisé et non voisé) l'on trouve une interface séparant [d] de [t]. Mais si l'on expérimente sur une langue comme le Thaï où il existe trois possibilités de voisement, on trouve deux interfaces séparant respectivement [d] de [t] et [t] de [t^h] ([t] aspiré). Des expériences de ce type montrent que la classification phonologique d'une langue résulte des systèmes d'interfaces K_s induits dans les espaces d'indices acoustiques par la perception catégorielle. Ces systèmes dépendent de la langue (l'interface [b]/[p] sur l'axe du VOT se trouve par exemple à 37 ms en anglais et à 5 ms en français) ce qui confirme la relativité des catégories phonologiques. Mais il semble que ces systèmes K_s faisant partie de l'état final stabilisé ('steady state') du langage proviennent tous d'un état initial K_0 ('initial state') de nature innée². En effet, un certain nombre d'expériences concordantes effectuées sur des enfants préverbaux semblent montrer qu'il existe une catégorisation universelle (génétiquement déterminée) du VOT définie par deux interfaces, l'une à environ -30, -20 ms et l'autre à environ +20, +30 ms.

L'existence d'une composante sensorielle innée est d'ailleurs, bien qu'elle remette en cause la vieille idée de la 'tabula rasa', nécessaire *a priori* à l'apprentissage de la langue. Car, ainsi que l'a remarqué P.D. Eimas (1980), pour qu'un enfant puisse faire l'apprentissage de sa langue maternelle, il faut qu'il puisse :

- i) discriminer de petites différences dans le signal acoustique ;
- ii) catégoriser une suite continue de valeurs acoustiques ;
- iii) apprécier l'organisation inhérente des unités de base, phonèmes et syllabes ;

- iv) repérer l'invariance de la perception malgré les variations du signal ;
- v) traiter l'information acoustique critique pour les distinctions phonétiques de façon dépendante du contexte.

2 LES INTERPRÉTATIONS EN CONFLIT

Il existe plusieurs interprétations de la perception catégorielle se distribuant selon les oppositions classiques sensoriel/cognitif et réductionniste/structural. Leur conflit est d'un intérêt épistémologique particulier dans la mesure où il montre que, faute d'une réflexion conceptuelle et théorique approfondie, il est pratiquement impossible de départager expérimentalement les paradigmes en présence.

2.1 *L'hypothèse sensorielle.*

Un certain nombre d'auteurs ont fait l'hypothèse que la perception catégorielle est une propriété générale de la perception, déductible de principes psychophysiques. Ainsi, selon J. Miller (Miller *et al.* 1976), elle ne ferait qu'indiquer la présence d'un seuil sensoriel masqué. Dans toutes les expériences, on considère un continuum associé à la variation d'un indice I et discrétisé en pas ΔI égaux. Au-dessous du seuil les variations ΔI ne seraient pas décelables (et donc non discriminables). Et au-delà du seuil, elles le redeviendraient en vertu de la loi de Weber-Fechner.

Cette hypothèse a été reprise par R. Pastore avec le souci de rendre compte de la perception catégorielle sans pour autant subordonner la discrimination sensorielle à l'identification cognitive. Selon Pastore, l'origine de la perception catégorielle est à trouver dans la structure des processus neuroperceptifs périphériques et relève d'une limitation soit interne soit externe qui est à la fois stable et définie plus précisément que le seuil de différenciation du continuum investigué. La limitation est interne quand elle correspond à un seuil masqué. Elle est externe quand elle fait intervenir un stimulus de référence auquel les stimuli sont comparés (Pastore *et al.* 1977.)

2.2 *L'hypothèse réductionniste des détecteurs de traits.*

Les expériences sur les enfants préverbaux remettent en cause les théories motrices classiques de la perception phonétique faisant de celle-ci un processus articulatoirement finalisé i. e. s'effectuant par 'reconstruction mentale' du processus articulatoire². En effet, les bébés percevant avant d'articuler, il faudrait supposer une détermination génétique du *feed back* de la perception sur l'articulation. Or 'to attribute this knowledge to the infant's biological endowment would seem to extend considerably the cognitive competencies that we are willing to impute to genetically determined factors' (Eimas 1974 : 518). D'autre part, certaines expériences d'écoute dichotique semblent indiquer que la perception s'effectue sur la base d'une recombinaison au niveau central de traits distinctifs extraits indépendamment les uns des autres du signal acoustique au niveau périphérique. Enfin, depuis leur découverte par Eimas et Corbitt en 1973, on dispose d'un certain nombre d'expériences concordantes sur les phénomènes d'*adaptation sélective* (Eimas et Corbitt, 1973).

Dans les expériences d'adaptation sélective, on considère un continuum W (par exemple [ba]–[pa]) et l'on présente répétitivement un des stimuli extrémaux aux sujets. Si, après cette adaptation, l'on construit l'interface K catégorisant W on constate que, relativement à la situation sans adaptation, elle s'est déplacée vers le stimulus adaptateur. Si maintenant on adapte non plus par l'intermédiaire d'un stimulus de la série testée mais par un stimulus ne comportant qu'un trait commun avec lui, on constate un déplacement identique de K (sélectivité de l'adaptation).

L'ensemble de ces résultats a conduit Eimas et Corbitt à faire l'hypothèse de *détecteurs de traits* qui sont des récepteurs neuro-sensoriels répondant sélectivement à des domaines précis de valeurs d'indices acoustiques. "Feature detectors" can be broadly defined as organizational configurations of the sensory nervous system that are highly sensitive to certain parameters of complex stimuli' (Abbs et Sussman 1971 : 24). Si l'on suppose qu'ils font partie de l'état initial phonétique de l'organisme, qu'ils ne sont sensibles qu'à des domaines bien limités de variation d'indices acoustiques et que leur domaine de réponse est acquis par interaction avec l'environnement, alors on peut expliquer simplement les expériences précédentes. Il suffit pour cela de supposer :

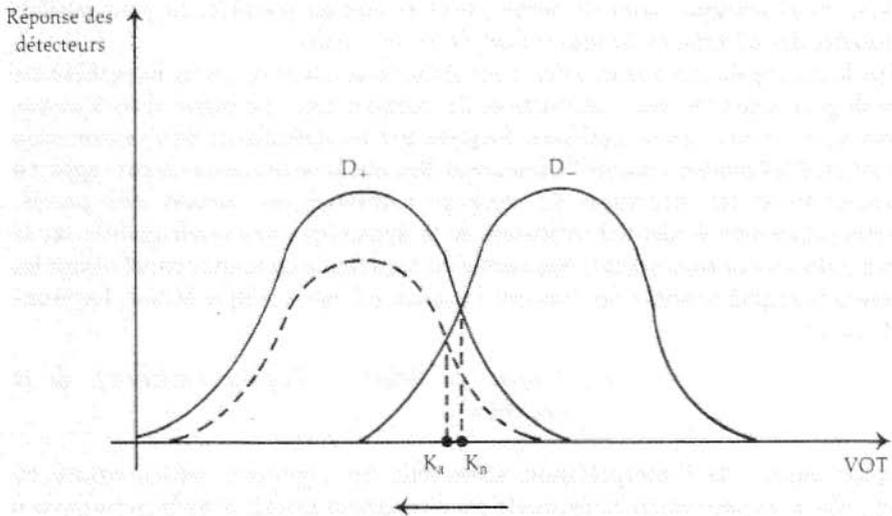
- i) que l'axe du VOT est recouvert par les domaines de deux (ou trois) détecteurs périphériques ;
- ii) que les réponses de ces détecteurs sont en compétition et que les dispositifs centraux de traitement de l'information ne sont sensibles qu'à l'excitation du détecteur donnant la réponse maximale ;
- iii) que la présentation répétée d'un même trait 'fatigue' le détecteur correspondant (cf. Fig. 5).

Si l'hypothèse réductionniste des détecteurs de traits a eu un grand impact c'est non seulement parce qu'elle fournit une explication simple de la perception catégorielle et des phénomènes d'adaptation sélective mais aussi parce qu'elle applique au domaine phonétique ce qui, depuis les travaux de Hubel et Wiesel, est bien connu pour la perception visuelle, et permet d'accéder à une conception neuro-physiologique unifiée de la perception. Mais, précisément parce qu'elle est de nature réductionniste, elle ne va pas sans faire théoriquement problème.

2.3 Critique de l'hypothèse des détecteurs de traits.

L'hypothèse des détecteurs de traits a fait l'objet d'un certain nombre de critiques non seulement de la part des tenants des théories motrices auxquelles elle s'oppose résolument mais aussi de la part de tenants de théories également réductionnistes. D'abord, ainsi que l'a remarqué R.L. Diehl, on peut expliquer les expériences d'adaptation sélective à partir de principes perceptifs d'une autre nature comme le principe de contraste selon lequel la perception privilégie la différence par rapport à l'identité. Si l'adaptation joue comme référence, les stimuli du continuum testé seront perçus 'plus tôt' comme différents et l'interface sera donc déplacée vers l'adaptateur

FIGURE 5 : Interprétation des déplacements d'interfaces en termes de détecteurs. L'axe du VOT est recouvert par le domaine de deux détecteurs D_+ (voisé) et D_- (non voisé). L'interface normale K_n correspond à la valeur du VOT pour laquelle il y a égalité des réponses de D_+ et de D_- . L'adaptation par un stimulus voisé 'fatigue' D_+ dont la réponse décroît (courbe de réponse en pointillés) ce qui déplace K_n vers K_a .



(Diehl *et al.*, 1978). D'autre part, ainsi que l'a noté J.S. Bryant, les effets d'adaptation sélective montrent simplement que les traits distinctifs phonologiques sont représentés neurologiquement. 'This does not necessarily mean that the sum of such representations is functional in perception as the feature detector notion implies. Rather, the cell or cells may simply be responding as a small part in a large pattern of neural response to the stimulus' (Bryant, 1978).

Quant à nous, nous adresserons à cette hypothèse des détecteurs une triple critique. D'abord, dans la mesure où il n'existe pas de rapports simples entre les indices acoustiques et les phonèmes encodés, elle conduit nécessairement à une prolifération des détecteurs. Ensuite, l'interface induite par la perception catégorielle sur un axe comme celui du VOT varie en fonction d'autres indices (par exemple la place d'articulation). Or cela est incompatible avec l'hypothèse de détecteurs de traits indépendants. Les espaces d'indices W naturellement associés aux gestalts que sont les phonèmes sont multidimensionnels et c'est de la morphologie des systèmes d'interfaces K qui les catégorisent qu'il faut rendre compte. En dimension 1 la véritable difficulté théorique posée par l'engendrement des K n'apparaît pas car, quel que soit le mécanisme générateur imaginé, les K se réduisent à des points isolés. Mais il n'en va plus du tout de même pour les dimensions supérieures où la considération de la morphologie des interfaces fournit un critère simple de falsification de l'existence de détecteurs. Supposons en effet qu'un espace d'indices W bidimensionnel soit recouvert par les domaines de réponse d'un nombre fini de détecteurs dont les surfaces de réponse sont, en analogie avec le cas unidimensionnel, des surfaces 'en cloche'. Alors K devrait être la projection sur W des courbes d'intersection de ces surfaces.

Mais les morphologies ainsi obtenues sont très différentes des morphologies génériques d'interfaces que l'on rencontre dans les phénomènes de transitions de phases et, plus généralement, dans les phénomènes critiques. Si donc l'on peut montrer (et il serait urgent de faire les expériences) que les morphologies observables d'interfaces phonétiques sont du second type et non du premier, on peut falsifier l'hypothèse des détecteurs du moins dans sa version naïve.

Mais la principale critique que l'on peut selon nous adresser à cette hypothèse est d'être dogmatiquement réductionniste et de marquer une régression théorique par rapport aux avancées structuralistes. Inspirée par les spécialistes de l'*information processing*, elle favorise, comme l'a remarqué Bryant, la reconnaissance par rapport à la perception et les structures de stockage mémoriel par rapport aux procès. Considérant comme évidente l'hypostase de la dynamique neuronale globale sur la structure du réseau sous-jacent, elle relève du paradigme computationnel visant les processus perceptifs comme des 'calculs' logiques (cf. par exemple Miller, Johnson-Laird, 1976).

2.4 *L'enjeu du débat et 'l'aporie fondatrice' de la phonétique.*

Le débat autour de l'interprétation sensorielle ou cognitive, réductionniste ou structurale de la perception catégorielle est d'un grand intérêt dans la mesure où il représente l'état actuel d'une difficulté que l'on pourrait appeler 'l'aporie fondatrice' de la phonétique et qui concerne la possibilité de comprendre sur des bases psychophysiques les phonèmes comme unités abstraites, linguistiquement fonctionnelles, définies non pas par des propriétés intrinsèques mais par un réseau de différences. Depuis Saussure et Jakobson, les phonèmes sont en effet visés comme des identités purement relationnelles (et non substantielles), comme des *gestalts* relevant d'un niveau formel de réalité supposé être ontologiquement autonome. Nous sommes donc en présence de deux catégorialités :

- i) Côté phonologie, une catégorialité structurale (jakobsonienne-hjelmslevienne) que l'on croit être de nature logico-combinatoire : identité de position, unité relationnelle, différence, discrimination, détermination réciproque, stratification, etc.
- ii) Côté phonétique, une catégorialité psychophysique : formes spectrales, déformations de ces formes, contrôle par des indices acoustiques, invariance/variabilité, catégorisation, interfaces, etc.

Et le problème est évidemment d'unifier – de 'dialectiser' – ces deux catégorialités. Or cela est impossible si l'on interprète la première de façon logico-combinatoire (i.e. formelle au sens formaliste du terme) et la seconde de façon réductionniste. D'où une véritable antinomie dialectique qui n'est pas sans rappeler les antinomies kantienne.

L'hypothèse des détecteurs de traits résout 'dogmatiquement' cette antinomie en annulant la catégorialité structurale au profit de la catégorialité psychophysique dont la valeur objective ne fait pas problème. Mais cela est théoriquement insatisfaisant car

tout le problème est précisément de constituer (et non pas de dénier *a priori*) la valeur objective de la catégorialité structurale. Il s'agit là d'un problème mathématique de 'schématisation' (en un sens kantien actualisé) de catégories conceptuelles qui n'est pas du tout de même nature que celui d'une modélisation.

S'il est donc légitime de postuler qu'à tous les niveaux de la hiérarchie et pour toute catégorie d'éléments du langage il faut supposer un aspect abstrait et fonctionnel, décrit uniquement en termes relationnels et un aspect substantiel qui se décrit selon le choix et le but du descriptiviste en termes d'articulation, de structure acoustique et de perception auditive' (Malmberg, 1974 : 210) encore faut-il comprendre le rapport de dépendance réciproque entre la forme et la substance phonétiques. En l'absence d'une telle compréhension, l'on ne peut opposer aux conceptions réductionnistes notoirement insuffisantes qu'un réalisme structural tout aussi insuffisant dans la mesure où, naturalisant des artefacts formels, il exige en dernière instance le recours à d'improbables hypothèse innéistes.

2.5 *L'a priori de la catégorisation paradigmatique.*

Pour comprendre le rapport de dépendance réciproque entre la forme et la substance de l'expression, il faut arriver à définir le contenu mathématique général de la situation abstraite générale exposée au §1.3 et montrer :

- i) que ce contenu permet de schématiser en les unifiant la catégorialité structurale et la catégorialité psychophysique qui en sont constitutives, et
- ii) que, par spécification de son contenu général, cette schématisation peut se spécialiser en modèles concordant avec les données expérimentales.

Autrement dit, avant de chercher à élaborer des modèles, il faut arriver à formuler mathématiquement l'*a priori* de la perception catégorielle c'est-à-dire l'*a priori* de la catégorisation paradigmatique.

Pour cela l'idée directrice est que le concept de catégorisation est la synthèse du concept (ancien) de classification et du concept (moderne) de contrôle. *A priori*, la situation abstraite générale exposée au §1.3 est du type suivant. On se donne un espace W d'indices acoustiques (que l'on supposera réduits à de simples paramètres) u_1, \dots, u_r . Un point $w = (u_1, \dots, u_r) \in W$ est donc un stimulus jouant le rôle d'entrée pour une 'boîte noire' perceptive S . Dans cette 'boîte noire' inobservable, un processus dynamique global X (que l'on peut supposer être réductible à un déterminisme neurophysiologique) définit des états internes A, B, C, \dots (des 'images acoustiques') et se trouve contrôlé par W . Qui plus est, une instance I sélectionne, $w \in W$ étant donné, l'état interne actuel, les autres états internes se trouvant virtualisés.

La situation générale est donc caractérisée :

- i) par un champ X_w de processus dynamiques contrôlés par W c'est-à-dire par une application $\sigma : W \rightarrow \mathcal{X}$ associant à chaque point $w \in W$ un 'point' X_w dans l'espace généralisé (l'espace fonctionnel) \mathcal{X} des processus internes considérés ;
- ii) par l'instance de sélection (d'actualisation) I .

Pour comprendre comment un tel système $S = (W, \mathcal{X}, \sigma, I)$ peut catégoriser son espace de contrôle il suffit alors de postuler que la perception n'est pas déterminée par la forme exacte de l'état interne A actuel mais seulement par son type qualitatif $\tau(A)$. Ce type qualitatif sera en général défini par l'action d'un groupe G sur l'espace fonctionnel \mathcal{X} . Si X est un point de \mathcal{X} , son orbite \bar{X} relative à l'action de G est constituée de processus qualitativement équivalents définissant des états internes de même type qualitatif que l'on pourra chercher à caractériser par les valeurs d'un système d'invariants τ_1, \dots, τ_k c'est-à-dire par des 'propriétés' des percepts (des phonèmes) associés.

Sur l'espace \mathcal{X} il existera en général une (et même des) topologie(s) permettant de dire quand deux processus $X_1, X_2 \in \mathcal{X}$ sont 'voisins'. Or l'existence d'une telle topologie et de l'action de G suffit à définir la stabilité structurelle des éléments de \mathcal{X} . Si en effet $X \in \mathcal{X}$, on dira que X est structurellement stable si tout $Y \in \mathcal{X}$ assez voisin de X est G -équivalent à X . Soit alors $K_{\mathcal{X}}$ le sous-ensemble de \mathcal{X} constitué des processus structurellement instables. $K_{\mathcal{X}}$ est une morphologie discriminante, inhérente à \mathcal{X} , qui le catégorise et classe les types qualitatifs de ses éléments stables. Autrement dit, tout espace \mathcal{X} de processus (et, plus généralement, tout espace de formes) où l'on sait définir les notions de déformation et de type qualitatif se trouve naturellement catégorisé par un sous-ensemble $K_{\mathcal{X}}$, appelé son ensemble catastrophique, qui géométrise le concept associé de classification. Autrement dit encore, interprétée mathématiquement (mais de façon très générale) en termes d'action de groupes sur des espaces généralisés et non plus de façon simplement logico-ensembliste, le concept classique de classification se traduit en un 'supplément' de géométrie et c'est ce supplément qui constitue *a priori* de la catégorisation paradigmatique.

Car soit $W \xrightarrow{\sigma} \mathcal{X}$ le champ exprimant le contrôle du système S par l'espace de contrôle W . σ 'plonge' en quelque sorte W dans \mathcal{X} et l'on peut supposer que ce 'plongement' est lui-même structurellement stable ce qui lui impose des contraintes drastiques en ce qui concerne sa complexité. Soit alors $K' = \sigma^{-1}(K_{\mathcal{X}} \cap \sigma(W))$ la trace de $K_{\mathcal{X}}$ sur W par l'intermédiaire de σ . La catégorisation de W induite par S et définie par le système de discontinuité K est déductible de K' lorsque l'on connaît l'instance de sélection I . En effet l'origine dynamique de K est la suivante. Soit A_w l'état actuel sélectionné par I en $w \in W$. Lorsque w varie dans W , A_w (supposé stable) varie à type qualitatif constant ce qui implique l'invariance du percept associé et des propriétés τ_1, \dots, τ_k . Mais il existera en général des valeurs critiques w_i de w pour lesquelles A_w entre en conflit (relativement à I) avec un autre état interne B_w et se trouve supplanté par lui après la traversée de w_i . K est donc constitué de ces valeurs critiques pour lesquelles il y a transition catastrophique de l'état actuel. Or ces catastrophes seront corrélées de façon réglée par I , soit à une déstabilisation interne de A_w (catastrophes dites de bifurcation) soit au fait que A_w entre en conflit avec B_w pour des raisons intrinsèques (catastrophes dites de conflit). Mais, pour un processus X , une déstabilisation d'un état interne associé ou un conflit entre deux états internes sont des causes d'instabilité et w appartiendra à K si et seulement si la situation en w est corrélée de la façon réglée par I à une situation appartenant à K' . C'est en ce sens que K est déductible de K' .

D'où l'*a priori* de la catégorisation paradigmatique : la catégorisation est la trace sur l'espace de contrôle des instabilités et des conflits des états internes qu'il contrôle.

2.6 *Éléments d'une critique structurale.*

Le dégagement de l'*a priori* (catastrophique non par hasard mais par essence) de la catégorisation, permet de développer dans un style assez proche du style kantien les 'Éléments' (Esthétique, Analytique des Concepts et Analytique des Principes) d'une 'Critique structurale' (Petitot, 1982). Donnons quelques indications.

2.6.1 *Traits internes discontinus et traits externes continus.*

À partir du moment où des indices (u_1, \dots, u_r) contrôlent des propriétés de percepts (τ_1, \dots, τ_k) qui sont des invariants de leur type qualitatif, la notion de trait distinctif doit être révisée. Il faut en effet distinguer les traits externes que sont les indices et les traits internes que sont les invariants qualitatifs. Les premiers varient de façon continue alors que les seconds varient au contraire de façon discrète et cela nécessairement selon les deux types phénoménologiques d'oppositions dégagées par Jakobson :

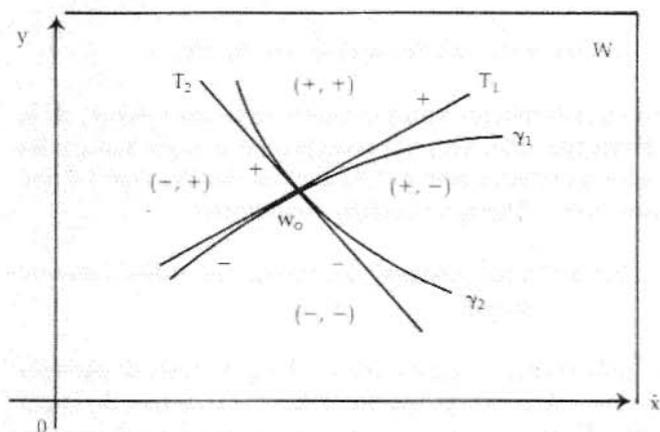
- i) les oppositions *qualitatives* correspondant aux catastrophes de conflit (compétition de deux invariants), et
- ii) les oppositions *privatives* correspondant aux catastrophes de bifurcation (présence/absence d'un même invariant).

Dans le cas du voisement par exemple, le trait externe est fourni par l'indice du VOT catégorisé par la perception catégorielle et le trait interne est fourni par l'opposition privative voisé/non voisé.

Dans le nouveau cadre théorique fondé par l'*a priori*, on peut donc enfin dépasser l'antinomie entre le continu et le discontinu et légitimer le binarisme jakobsonien sur la base même de la variation continue des indices. Cela résout une des difficultés majeures des descriptions phonétiques : 'one of the major difficulties in achieving this kind of description is in relating the essentially continuous nature of speech with the essentially discontinuous nature of linguistic description' (Ladefoged 1972: 276, cf. aussi Massaro, 1972).

Ceci dit, le problème n'est pas pour autant complètement résolu et cela pour deux raisons. D'abord les traits externes ne coïncident pas en général avec les indices. Soit en effet W un espace d'indices (u_1, \dots, u_r). L'essentiel de l'information est fournie par la morphologie discriminante K (l'ensemble catastrophique) induite dans W par la 'boîte noire' perceptive. Or il n'existe aucune raison *a priori* pour que le repère R de W constitué par les axes u_1, \dots, u_r soit 'adapté' à K . La position de K par rapport à R constitue même la donnée fondamentale sur la non indépendance (et donc sur l'intégration) des indices. Il est par suite naturel de chercher un repère R_K de W adapté

FIGURE 6 : Discrétisation d'un repère adapté à $K (= \gamma_1 \cap \gamma_2)$. Chaque région de W définie par K est 'codable' par un doublet (\pm, \pm) .

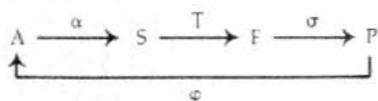


à K . C'est à ce repère que correspondent les traits externes. Cela permet d'ailleurs de comprendre pourquoi ceux-ci peuvent eux aussi, bien que continus, être en général discrétisés et réduits à une opposition de type $+/-$ (qu'il ne faut pas confondre avec les oppositions qualitatives ou privatives décrivant les traits internes). En effet il ne s'agit pas tant de repérer les points de W que les domaines de W que K différencie, classifie, positionne les uns par rapport aux autres. Or cela peut en général se faire avec une information discrète. Supposons par exemple que W soit de dimension 2 (indices x et y) et que K soit constitué de deux courbes γ_1 et γ_2 s'intersectant transversalement en w_0 . K partage donc W en quatre domaines et pour repérer ceux-ci localement il suffit de considérer le repère adapté constitué des tangentes respectives T_1 et T_2 à γ_1 et γ_2 en w_0 , chaque domaine devenant 'codé' par une demi-tangente (+ ou -) à γ_1 et une demi-tangente (+ ou -) à γ_2 (cf. Fig. 6).

Il s'ensuit qu'en ce qui concerne les traits externes, le binarisme jakobsonien (*a priori* valide pour les traits internes) doit être interprété comme un renseignement sur le type de complexité locale des interfaces phonétiques : celles-ci sont localement assez simples pour que leurs repères adaptés soient discrétisables de façon binaire. Et le fait que la combinatoire des traits distinctifs soit une combinatoire contrainte ainsi que le phénomène de stratification des traits (i.e. les relations de marquage et de dominance) sont autant de renseignements sur ce qui sépare la structure locale de K de celle décrite à la Fig. 6 (qui est un cas de combinatoire libre) ainsi que sur sa structure globale.

La seconde raison faisant que des modèles catastrophiques élémentaires ne peuvent jouer pour la perception catégorielle que comme *a priori* et non pas comme modèles exacts est que le système phonétique est un système à deux niveaux de contrôle. A un premier niveau, un espace A de paramètres articulatoires contrôle des spectres continus appartenant à un espace S . On a donc un premier champ $\alpha : A \rightarrow S$ possédant l'avantage éminent d'être observable et correspondant à la variation des spectrogrammes en fonction de l'articulation (cf. les travaux classiques de P. Delattre et du

Laboratoire Haskins). Mais à un second niveau, les transformés auditifs des spectres acoustiques forment un espace F contrôlant un espace P de processus psychiques définissant des percepts. On a donc un second champ $\sigma : F \rightarrow P$ inobservable directement et dont l'espace de contrôle n'est pas un espace de paramètres mais lui-même un espace de formes. Les deux niveaux sont reliés d'une part par la transformation auditive $T : S \rightarrow F$ et d'autre part par un *feed-back* $\varphi : P \rightarrow A$ exprimant que la perception se fait par reconstruction d'un schème moteur articuloire et que, corrélativement, l'articulation est finalisée par un espace de cibles phonémiques internalisé. Dans ce cadre général :



Les traits distinctifs peuvent être conçus :

- i) comme des repères adaptés aux interfaces induites dans A soit par S , soit par F , soit par P , i.e. comme des traits externes de nature articuloire qui seraient la trace d'instabilités et de conflits d'invariants des spectres acoustiques, de leurs transformés auditifs ou de leurs images acoustiques ;
- ii) comme des propriétés des types qualitatifs des spectres acoustiques, i.e. comme des traits internes de nature acoustico-auditive caractérisant des éléments de S ou de F ;
- iii) comme des repères adaptés aux interfaces induites dans S ou F par la perception P , i.e. comme des traits externes de nature acoustico-auditive qui seraient la trace d'instabilités ou de conflits 'd'attracteurs' de dynamiques neurologiques (corticales) inobservables et définies sur des espaces de configuration de très grande dimension ; et enfin
- iv) comme des invariants de ces 'attracteurs', i.e. comme des traits internes de nature perceptive.

2.6.2 La condition de discrétisation.

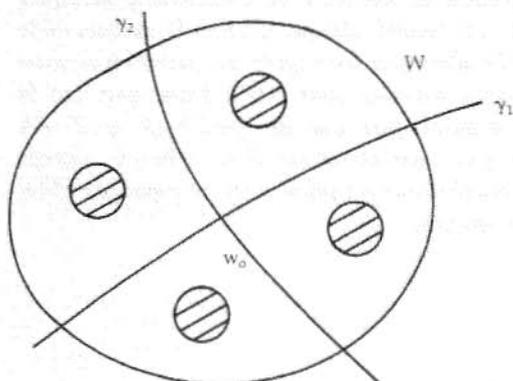
L'*a priori* de la catégorisation permet donc de dépasser de façon fondée l'antinomie du discret et du continu et de concilier les deux conceptions que l'on peut se faire d'un phonème :

- i) une gestalt servant de prototype à une classe d'allophones (critère logico-combinatoire de l'identité d'un phonème),
- ii) un domaine dans un espace de contrôle (critère positionnel-structural de l'identité d'un phonème).

Les phonèmes comme gestalts prototypiques ne sont que les 'capitales' des domaines définis par la catégorisation K de W .

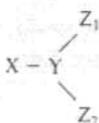
Mais l'*a priori* fait, qui plus est, apparaître un conflit irréductible entre le critère

FIGURE 7 : La condition de discrétisation. Il existe une correspondance biunivoque entre les 'capitales' et les composantes connexes de $W - K$ ($K = \gamma_1 \cup \gamma_2$).

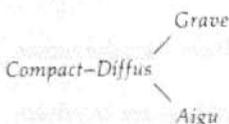


logico-combinatoire et le critère positionnel-structural de l'identité. En effet, pour que ces critères soient compatibles il faut que soit satisfaite une condition que nous appellerons la condition de discrétisation. Cette condition dit qu'il existe une correspondance biunivoque entre les prototypes et les composantes connexes du complémentaire de K dans W . Elle est satisfaite pour la catégorisation représentée à la Figure 6 (cf. Fig. 7)

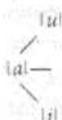
Mais elle n'a aucune raison d'être satisfaite *a priori*. Considérons par exemple dans un espace de contrôle W de dimension 2, une interface K s'arrêtant en un point δ . K est un seuil séparant deux déterminations A et B par une opposition qualitative A/B . Mais en δ le seuil disparaît et le terme neutre-complexe $A*B$ de l'opposition se trouve engendré. Il existe donc trois prototypes pour une seule composante connexe. Or ce type d'interfaces est inéliminable de la théorie. Dans le cas physique des transitions de phases (qui, nous l'avons vu, devrait devenir paradigmatique pour la compréhension et l'explication des phénomènes de perception catégorielle), il correspond à ce phénomène fondamental qu'est l'existence de points critiques (arrêt de l'interface liquide/gaz). Et en phonétique il correspond très exactement, par choix d'un repère et par discrétisation, aux descriptions en termes de traits distinctifs de type



c'est-à-dire à la subordination d'une opposition Z_1/Z_2 à l'un des termes Y d'une opposition dominante X/Y . Mais tel est le cas en particulier de la hiérarchie élémentaire



qui classifie le triangle universel des voyelles cardinales



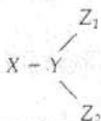
(cf. Fig. 8)

Il s'agit là d'un phénomène général et fondamental. la détermination d'identités positionnelles par des systèmes de différences peut entrer en conflit avec les unités discrètes associées et violer le principe logique d'identité. Les identités positionnelles ne correspondent pas nécessairement à des identités isolables et à des unités objectales discrètes. En ce sens, l'*a priori* de la catégorisation légitime en droit le principe structural (saussurien-jakobsonien) du primat ontologique de la différence sur l'identité et montre que le contenu éidétique des catégories structurales ne peut être de nature logico-combinatoire.

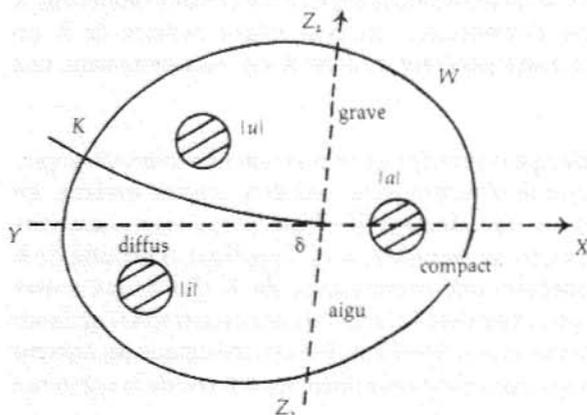
2.6.3 Le principe de relèvement phénoménologique.

À un niveau plus profond et plus général, l'*a priori* catastrophique de la catégorisation permet de concilier les deux conceptions jusqu'ici rivales de la phonétique, la

FIGURE 8 : Cas où la condition de discrétisation n'est pas remplie. K est une interface s'arrêtant au 'point critique' δ (disparition du seuil séparant les deux régimes). Le 'repère adapté' est constitué par l'axe dominant tangent à K en δ , séparé en deux zones : Y (+ par exemple) correspondant à K et X correspondant à la zone opposée. A cet axe se subordonne un autre axe Z_1/Z_2 repérant les deux zones séparées par K du côté Y. Le repère est donc du type



Il existe trois 'capitales', une dans la zone X codée par $(-, 0)$ où 0 symbolise la neutralité relativement à l'opposition subordonnée Z_1/Z_2 , et deux dans la zone conflictuelle codées respectivement par $(+, -)$ et $(+, +)$. Mais il n'existe qu'une composante connexe de $W - K$.



conception (réductionniste) 'substance based' et la conception (structurale) 'form based'. L'idée directrice (idée que ce que nous venons de présenter aura, nous l'espérons, rendue évidente, et d'une évidence apodictique) est que *la forme relationnelle de l'expression équivaut à la phénoménologie de la substance de l'expression*. Dans cette affirmation c'est le terme de 'phénoménologie' qui est essentiel. Ainsi que nous l'avons noté en 2.4., tant que l'on conçoit la substance de façon réductionniste (psycho-physique) et la forme de façon formaliste il est impossible d'arriver à une théorie phonético-phonologique unitaire. Pour y arriver il faut fonder la théorie sur l'*a priori* de la perception catégorielle, c'est-à-dire :

- i) concevoir la forme relationnelle comme une description logico-combinatoire des morphologies discriminantes (W, K) i.e. de la pure phénoménologie des catégorisations ; et
- ii) supposer à titre purement implicite un mécanisme générateur X (un processus psychique inobservable) susceptible d'engendrer ces morphologies.

Cette seconde hypothèse pourra certes paraître difficile à admettre. Elle est pourtant inévitable dans la mesure où elle ne fait que formuler ce truisme que le psychisme est une 'boîte noire'. Le problème n'est pas celui de sa pertinence mais celui de sa portée opératoire. Le réductionnisme se caractérise par la croyance optimiste que les progrès de la neurophysiologie permettront un jour d'expliquer causalement le mécanisme X. Mais il s'agit là d'une naïveté. En effet – et c'est là que la théorie des catastrophes devient épistémologiquement et méthodologiquement primordiale – une analyse mathématique poussée des situations de contrôle $\sigma : W \rightarrow \mathcal{X}$ constituant l'*a priori* de la catégorisation montre

- i) que les mécanismes générateurs X susceptibles d'engendrer une morphologie (W, K) sont hautement surdéterminés par rapport à la géométrie de K ;
- ii) qu'en ce sens les morphologies sont en grande partie indépendantes des substrats et soumises à des contraintes purement géométriques de nature formelle et non matérielle³ ; et enfin
- iii) qu'il est possible de 'remonter' de la phénoménologie de K à des contraintes sur X et que, dans les cas les plus élémentaires, on peut même déduire de K un modèle minimal X_0 de mécanisme générateur, dont X est nécessairement une complexification.

Nous appellerons ce dernier principe *le principe de relèvement phénoménologique*. Il inverse l'ordre de déduction que le réductionnisme considère comme évident. En effet au lieu de chercher à déduire causalement K d'une connaissance explicite (neurophysiologique) de X on cherche au contraire, si on l'applique, à déduire de K (i.e. de la manifestation phénoménale) une connaissance du X réel donné à titre purement implicite. Ces deux acquis essentiels de la théorie des catastrophes que sont l'indépendance par rapport au substrat et le relèvement phénoménologique permettent de comprendre comment, en tant qu'équivalente à la phénoménologie de la substance

de l'expression, la forme relationnelle de l'expression peut être ontologiquement autonome.

2.7 Les interprétations précatastrophiques de la perception catégorielle.

En tentant d'abstraire le contenu éidétique de la perception catégorielle, un certain nombre d'auteurs en sont arrivés à des conceptions assez proches de son *a priori* catastrophique.

2.7.1 Les modèles de prototypes.

Étant donnée une catégorisation phonétique (W, K), les allophones correspondant au voisinage de K ne sont pas en général naturellement productibles et doivent être synthétisés. Les allophones naturellement produits (articulables) s'accumulent dans des zones (des sortes de 'capitales') correspondant à des valeurs typiques des indices acoustiques. En effet, comme le rappelle Lindblom, si la production articuloire de la parole est bien un procès 'output-oriented' i.e. finalisé par des cibles perceptives, ces cibles sont choisies de façon à maximiser la stabilité acoustique par rapport à la variabilité articuloire. Il s'agit là d'un des caractères essentiels de ce que Lindblom a appelé la 'distinctiveness condition' conçue comme 'evolutionary conspiracy' (Lindblom, 1972). De même, Ladefoged a introduit la notion de 'conventions d'interprétation' associant à la forme idéale d'une unité linguistique discrète des valeurs typiques des paramètres de contrôle (Ladefoged 1972).

Selon ce point de vue que D. Massaro a qualifié de 'template matching scheme' (Massaro, 1972), les phonèmes sont donc conçus comme des gestalts unitaires. On admet qu'il existe, dans l'espace P des percepts (supposé muni d'une distance perceptive), un nombre fini de patrons phonétiques, de prototypes p_1, \dots, p_n . Étant donné un percept p , un dispositif cherche à optimiser la distance de p aux p_i et reconnaît p comme une occurrence du p_i (comme un allophone du phonème p_i) dont la distance à p est minimale. La catégorisation K est alors composée d'un ensemble de morceaux de médiatrices. Nous l'appellerons une T -classification. (cf. Fig. 9).

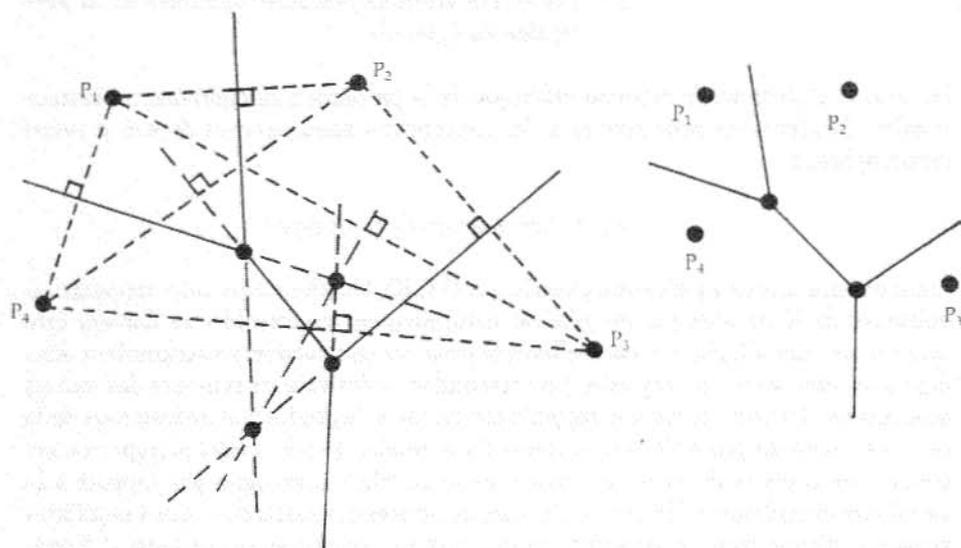
Un tel modèle a été développé par Repp à partir d'une réfutation de l'hypothèse réductionniste de la recombinaison des traits supposés détectés.

2.7.2 Les modèles précatastrophiques.

Parallèlement aux modèles de prototypes associés à des T -classifications on a développé (mais sans mathématiques adéquates) des modèles de nature proprement précatastrophique associés à des K -classifications. Dans ces modèles qui formulent l'*a priori* $\sigma : W \rightarrow \mathcal{X}$ de la catégorisation, on part des hypothèses suivantes :

- i) les spectres acoustiques continus (les spectres harmoniques n'étant pas, on le sait, phonologiquement pertinents) sont des formes déformables contrôlés ;

FIGURE 9 : Exemple de T-classification. Dans un 'espace' psychologique (internalisé) de cibles phonétiques, espace muni d'une 'distance' perceptive, on se donne des prototypes p_1, \dots, p_n . Le domaine d'un prototype p_i est constitué par l'ensemble des points p dont la distance à p_i est moindre que celle aux $p_j \neq i$. Le système d'interfaces est donc constitué par des segments de médiatrices.



- ii) l'audition extrait de l'information qualitative de ces spectres ;
- iii) certains types de spectres sont stables relativement à leur contrôle et d'autres non ;
- iv) la perception ne retient des types qualitatifs spectraux que des invariants ;
- v) les phénomènes de perception catégorielle sont dus à des transitions catastrophiques de ces types qualitatifs sous l'action du contrôle. Ainsi que l'a noté Pisoni, il peut y avoir perception catégorielle chaque fois que des entités complexes contrôlées par des paramètres comportent plusieurs qualités distinctives dont la présence ou l'absence définissent des domaines du contrôle, ces catégories étant étiquetables et codables dans la mémoire à court terme (Pisoni 1979).

La première idée directrice est donc qu'il faut traiter les percepts 'not as bundles of separately extracted phonetic features but as integral multidimensional entities whose dimensions are inseparable aspects of the whole pattern'. 'The dimensions are assumed to reflect the auditory properties of the stimulus and thus are continuous, not binary. Instead of representing speech sounds as matrices of discrete feature values, they are conceptualized as points in a continuous multidimensional perceptual space' (Repp *et al.* 1978). Il ne s'agit plus de postuler que la perception repose sur une détection d'indices qui seraient ensuite combinés en un percept unifié mais de partir de l'*a priori* (non combinatoire) du contrôle et d'imaginer que 'the sound simply initiates the unfolding of a complex pattern of neural response that directly supports the phenomenal experience' (Bryant 1978: 616). Ce point de vue structural (holiste)

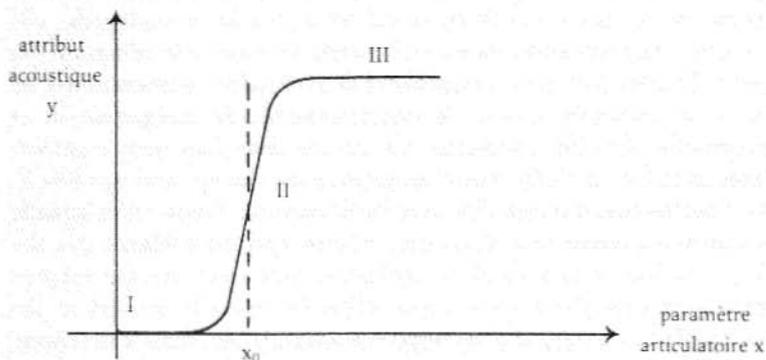
permet de comprendre la perception catégorielle sans pour autant subordonner la discrimination sensorielle à une identification cognitive. En effet, ce n'est que si l'on opte pour un modèle de prototypes que l'identification est cognitive puisqu'elle s'effectue par comparaison avec un matériel stocké dans la mémoire à long terme. Si l'on admet en revanche qu'elle consiste en la valeur d'invariants qualitatifs, elle équivaut alors à une discrimination catégorielle sans aucune subordination du sensoriel au cognitif. D'autre part, pour comprendre dans ce cadre l'existence chez les enfants préverbaux de capacités innées de discrimination, de catégorisation et d'invariance perceptuelle, il suffit d'admettre un cablage génétiquement contraint incluant dans l'état initial de la 'boîte noire' perceptive un champ inné $\sigma_0 : F \rightarrow P$. Cette hypothèse est parfaitement compatible avec l'affirmation d'Eimas selon laquelle 'the fact that constancy is present in prelinguistic infants is strong evidence that the means by which it is achieved is a result of biological constraints on the infant's perceptual system, such that the system must either be innately atuned to the invariance in the signal or innately able to impose constancy on acoustic diversity' (Eimas 1980). Mais elle n'implique en rien (comme elle le fait chez Eimas) la nécessité d'admettre l'existence de détecteurs. On peut penser plutôt avec Jusczyk que le champ initial $\sigma_0 : F \rightarrow P$ – qui serait un universel psychophysique de notre espèce issu phylogénétiquement de la structure de l'appareil auditif des mammifères supérieurs – détermine les propriétés de saillance perceptive des stimuli acoustiques auxquelles 's'accrocherait' le langage. 'As the child begins to acquire the phonological structure of the language, one would expect to see him weight the various acoustic cues present in the speech signal according to their salience in marking distinctive contrasts in the language' (Jusczyk, 1980).

Au cours de l'acquisition du langage, la K-classification universelle K_0 déterminée par σ_0 se trouverait alors déformée, complexifiée et spécifiée par l'environnement linguistique et progressivement subordonnée à une T-classification de nature cognitive.

Mais c'est sans conteste Kenneth Stevens qui s'est le plus approché de l'*a priori* catastrophique de la perception catégorielle, en particulier dans son article classique *The Quantal Nature of Speech* (Stevens, 1972). Stevens se situe au premier niveau de contrôle défini par le champ $A \xrightarrow{\alpha} S \xrightarrow{T} F$ et introduit deux idées directrices.

- i) La relation établie entre A et S (i.e. entre le niveau articulatoire et le niveau (audio) acoustique) par le contrôle α est 'non linéaire'. Il y a des zones de A où les spectres associés sont stables relativement à α et ces zones constituent la base du code phonologique. 'For a particular range of an articulatory parameter, the acoustic output from the vocal tract seems to have a distinctive attribute that is significantly different from the acoustic attributes for some other region of the articulatory parameter. Within this range of articulation, the acoustic attribute is relatively insensitive to perturbations in the position of the relevant articulatory structure' (Stevens, Perbell, 1977, cité dans Mac Neilage, 1979).
- ii) Les attributs acoustiques dont on étudie la stabilité et l'invariance relativement aux variations du contrôle articulatoire sont des configurations spectrales d'ordre

FIGURE 10 : Relation typique entre un paramètre articuloire x et un attribut acoustique y . Dans la zone I ($x < x_0 - \epsilon$) l'attribut y est absent. Dans la zone III ($x > x_0 + \epsilon$) l'attribut y est présent et insensible aux variations de x . Dans la zone II en revanche ($x_0 - \epsilon \leq x \leq x_0 + \epsilon$) il y a transition 'catastrophique' : y est instable relativement aux variations de x .



supérieur i.e. des propriétés qualitatives et globales des spectres et non pas des indices élémentaires isolés. Autrement dit, ce sont des propriétés de la 'gross shape' des spectres et, plus précisément, de leurs transformés auditifs (Stevens et Blumstein, 1978).

Ainsi, selon Stevens, le rapport entre un paramètre articuloire et un attribut acoustique qu'il contrôle est typiquement de nature catastrophique ainsi que l'atteste la Figure 10 reprise de Stevens (1972).

'There are certain conditions for which a small change in some parameter describing the articulation gives rise to an apparently large change in the acoustic characteristics of the output, there are other conditions for which substantial perturbations of certain aspects of the articulation produce negligible changes in the characteristics of the acoustic signal' (Stevens, 1972, 52). Cela permet de comprendre la base acoustico-articuloire de la forme phonologique en échappant aussi bien au réductionisme physique qu'au réalisme structural (cf. 2.6.3).

3 LES PRINCIPES DES MODÈLES CATASTROPHIQUES

La théorie des catastrophes se propose d'étudier *a priori* $\sigma : W \rightarrow \mathcal{X}$ lorsque les processus considérés sont définis par des systèmes dynamiques (des champs de vecteurs) X_w sur des variétés différentiables M . Ce programme étant trop vaste (la structure des systèmes dynamiques génériques constituant à elle seule un redoutable problème), elle se restreint en fait dans la théorie dite élémentaire aux systèmes dynamiques X_w dérivant de potentiels $f_w : M \rightarrow \mathbb{R}$ (\mathbb{R} étant la droite réelle). Il s'agit donc de comprendre, M étant une variété différentiable (supposée compacte) de dimension n , la structure de l'espace fonctionnel \mathcal{F} des applications différentiables $f : M \rightarrow \mathbb{R}$ et des champs $\sigma : W \rightarrow \mathcal{F}$ envoyant dans \mathcal{F} un contrôle W et cela en se

focalisant sur la géométrie des ensembles catastrophiques $K_{\mathcal{F}}$ et K . Donnons très brièvement quelques idées directrices.*

3.1 Le type qualitatif

Le type qualitatif d'un potentiel $f : M \rightarrow \mathbb{R}$ est assimilé à son type différentiable défini par l'action du groupe $G = G_M \times G_{\mathbb{R}}$ où G_M (resp. $G_{\mathbb{R}}$) est le groupe des difféomorphismes (des automorphismes pour la structure différentiable) de M (resp. de \mathbb{R}). Si $f \in \mathcal{F}$, son orbite f sous l'action de G est donc constituée des $g = \psi \cdot f \cdot \varphi^{-1}$ où $\varphi \in G_M$ et $\psi \in G_{\mathbb{R}}$.

Si $f \in \mathcal{F}$, l'information qualitative (G - invariante) essentielle sur sa structure est fournie par ses éléments critiques. Soit $x \in M$ et $D_f(x)$ l'application linéaire tangente de f en x envoyant linéairement l'espace vectoriel tangent $T_x M$ de M en x sur l'espace vectoriel tangent.

$T_{f(x)} \mathbb{R} \simeq \mathbb{R}$ de \mathbb{R} en $f(x)$. Si (x_1, \dots, x_n) est un système de coordonnées locales en x , la matrice de $D_f(x)$ est la matrice uniligne

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

On dit que x est un point critique de f si $D_f(x)$ n'est pas de rang maximal 1 en x , i.e. si toutes les dérivées partielles $\partial f / \partial x_i$ s'annulent en x . $y = f(x)$ s'appelle alors valeur critique. Si y est non critique (i.e. régulière) alors f est 'localement triviale' ce qui signifie :

- i) l'image inverse $f^{-1}(y)$ de y par f est une sous-variété M_y de M de codimension 1 (de dimension $n - 1$), et
- ii) pour un voisinage U de y , $M_U = f^{-1}(U)$ est difféomorphe au produit direct $M_y \times U$. (cf. Fig. 11).

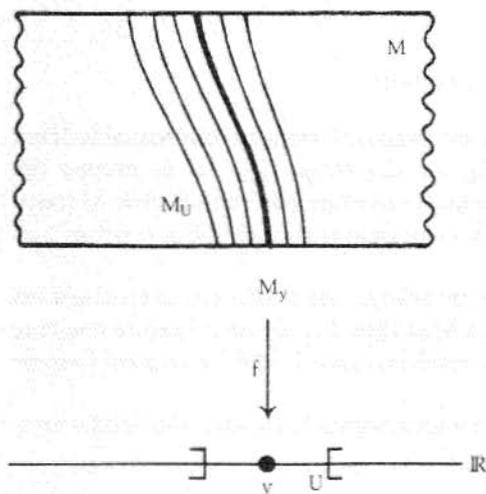
Lorsque x est critique, son 'degré' de criticité est mesuré par la plus ou moins grande dégénérescence des dérivées supérieures de f en x . La dérivée seconde de f est donnée par la matrice $n \times n$ (symétrique) des dérivées partielles secondes

$$H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right).$$

Si H est de rang maximum n , on dit que x est un point critique non dégénéré et l'on définit son indice comme l'index de la forme quadratique H . Les éléments critiques de f sont ses points critiques ainsi que leur 'degré' de criticité, ses valeurs critiques ainsi que leur multiplicité. Ce sont des invariants du type différentiable.

3.2 Le critère de stabilité structurelle

Un potentiel $f \in \mathcal{F}$ est structurellement stable si tout potentiel g assez voisin de f (pour la topologie différentiable) est G - équivalent à f , autrement dit si l'orbite f de f

FIGURE 11 : Trivialité locale d'un potentiel $f : M \rightarrow \mathbb{R}$ 

contient un voisinage entier de f . Un théorème fondamental, le théorème de Morse, affirme que, si M est compacte, f est structurellement stable si et seulement si :

- i) tous ses points critiques sont non dégénérés, et
- ii) toutes ses valeurs critiques sont distinctes.

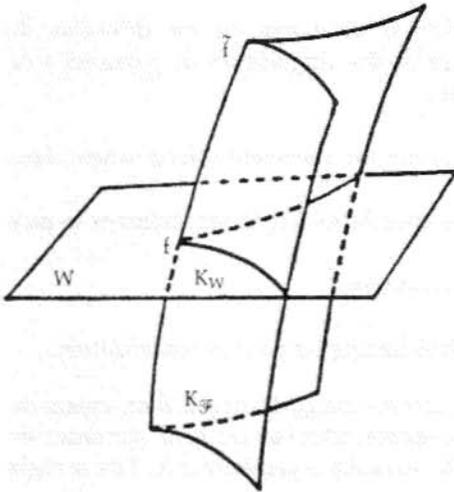
Il n'y a donc que deux causes possibles d'instabilité, la dégénérescence de points critiques (catastrophes de bifurcation) et l'égalité de valeurs critiques (catastrophes de conflit).

3.3 Généricité et transversalité.

L'équivalent du développement de Taylor d'une fonction différentiable $f : \mathbb{R} \rightarrow \mathbb{R}$ est ici donné par les applications $j^k f : M \rightarrow J^k$ qui associent à $x \in M$ la suite $j^k f(x)$ des dérivées jusqu'à l'ordre k de f en x (suites dites k -jets des applications et parcourant des fibrés J^k de base $M \times \mathbb{R}$). Or il est facile de voir que la première condition du théorème de Morse s'exprime en disant que le 1-jet $j^1 f(x)$ de f en x est transversal sur la section zéro de J^1 ($j^1 f(x) = 0$ exprime que x est critique et la transversalité exprime la non dégénérescence). La seconde condition s'exprime également en termes de transversalité. En vertu d'un théorème de R. Thom, cela implique que la stabilité structurelle est une propriété générique des potentiels i.e. que la non stabilité est 'exceptionnelle'.

3.4 Détermination finie, déploiements universels et modèles transverses.

Soit $K_{\mathcal{F}}$ l'ensemble catastrophique global, inhérent à \mathcal{F} , composé des potentiels structurellement instables. Si $f \in K_{\mathcal{F}}$, la géométrie de $K_{\mathcal{F}}$ au voisinage de f classe les

FIGURE 12 : La structure de \mathcal{F} au voisinage d'un potentiel f 'faiblement' instable.

types stables que l'on peut obtenir à partir de f par petites déformations. Si f est 'infinitement' instable (par exemple si f est constante au voisinage d'un point $x \in M$), la géométrie locale de K en f sera 'chaotique'. La théorie élémentaire s'intéresse aux cas de 'faible' instabilité où les conditions suivantes sont vérifiées.

- i) f est de détermination finie i.e. équivalente à l'un de ses jets d'ordre fini (qui est un polynôme). Cela implique que f soit de codimension finie c'est-à-dire que l'orbite \tilde{f} de f admette en f un supplémentaire W relativement à \mathcal{F} de dimension finie.
- ii) Au voisinage de f , la paire $(\mathcal{F}, K_{\mathcal{F}})$ est équivalente au produit direct de la paire (W, K) , où $K = K_{\mathcal{F}} \cap W$, par l'orbite \tilde{f} . (cf. Fig. 12).

Dans ce cas on dit que (W, K) est un modèle transverse de f . W étant de dimension finie k , il est isomorphe à un voisinage W' de l'origine de \mathbb{R}^k et l'on peut donc interpréter W comme un champ (un déploiement) $\sigma : W' \rightarrow \mathcal{F}$ qui à $w \in W'$ associe l'élément $f_w \in W$ correspondant. σ s'appelle le déploiement universel de f , f son centre organisateur. Le théorème du déploiement universel dit essentiellement que σ est unique à équivalence près et permet de reconstruire tous les déploiements de f .

3.5 Degrés d'instabilité et stratifications.

Si $f \in K_{\mathcal{F}}$ correspond à la situation élémentaire précédente, la possibilité de la stabiliser progressivement par paliers successifs se lira sur la géométrie de K (où (W, K) est un déploiement universel) : K sera une stratification c'est-à-dire un 'empilement' de lieux singuliers de dimensions décroissantes, chaque strate correspondant à un degré précis d'instabilité. Ces stratifications géométrisent le concept de classification.

3.6 *Formes normales et théorème de classification.*

Le théorème de Thom sur les catastrophes élémentaires est un théorème de classification (à équivalence différentiable près) des singularités de potentiel f de codimension ≤ 4 . Pour chaque cas, il fournit :

- i) le nombre de variables internes (1 ou 2) qui interviennent effectivement dans l'instabilité de f ;
- ii) une forme normale (polynomiale) de la singularité i.e. le représentant le plus simple de sa classe d'équivalence ;
- iii) une forme normale de son déploiement universel.

3.7 *Règles méthodologiques pour la modélisation.*

Étant donné un phénomène se manifestant comme catégorisation K d'un espace de contrôle W , la méthodologie catastrophique consiste, nous l'avons vu, à 'remonter' de cette phénoménologie à des contraintes sur le mécanisme générateur X . Elle se règle sur certains principes dont voici deux exemples.

- (a) La phénoménologie des catastrophes élémentaires étant connue on peut, si l'on rencontre empiriquement une de celles-ci, faire l'hypothèse que la catastrophe associée C gouverne le phénomène, en est une 'infrastructure' autrement dit que le mécanisme générateur réel X est une complexification de C qui est phénoménologiquement non pertinente.
- (b) D'autre part, étant donné un champ $\sigma : W \rightarrow \mathcal{X}$, on peut poser *a priori* que ce champ, dans la mesure où il existe concrètement, est nécessairement structurellement stable. Cette stabilité se traduit par une condition de transversalité de σ sur $K_{\mathcal{X}}$. Or la conséquence de cette transversalité *a priori* est que $\sigma(W)$ doit éviter les strates de $K_{\mathcal{X}}$ dont la codimension est supérieure à la dimension de W . Il s'agit là d'un principe fondamental : *la dimension de l'espace de contrôle W borne drastiquement la complexité des morphologies qui peuvent s'y déployer de façon structurellement stable.*

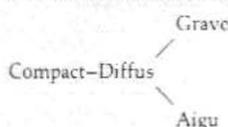
4 PERCEPTION CATÉGORIELLE ET CATASTROPHES ÉLÉMENTAIRES

La perception catégorielle transformant le flux phonétique en code phonologique étant un cas perceptif de phénomène critique, il est, nous l'avons vu, naturel de lui appliquer la méthodologie catastrophique. Pour l'instant ces méthodes ne peuvent pas être poussées très loin faute de résultats expérimentaux précis. Proposons toutefois quelques indications programmatiques.

4.1 *La classification des voyelles.*

Bien que perçues de façon continue à l'état isolé, les voyelles sont perçues de façon plutôt catégorielle lorsqu'elles sont prises dans un flux phonétique. Revenons à leur

description phonologique classique en termes de traits distinctifs :



Acoustiquement parlant, cette description se fait traditionnellement de la façon suivante :

- i) en tant qu'états stationnaires, les voyelles sont caractérisées par la fréquence et l'intensité de leurs formants i.e. par des paramètres quantitatifs ;
- ii) les deux premiers formants F_1 et F_2 sont essentiels et l'on peut s'y restreindre en première approximation ;
- iii) qualitativement parlant (i.e. au niveau de la perception phonologique des configurations spectrales d'ordre supérieur), la compacité de /a/ est due au regroupement de F_1 et F_2 dans la zone centrale du spectre ;
- iv) le caractère diffus de /i/ et de /u/ est dû au contraire à la séparation de F_1 , F_2 , le caractère aigu de /i/ étant dû à la dominance de F_2 sur F_1 et le caractère grave de /u/ à la dominance de F_1 sur F_2 .

L'on voit qu'une telle description qualitative dérive de la situation limite suivante :

- i) un espace de contrôle W de dimension 2 ayant pour coordonnées les fréquences ν_1 et ν_2 de F_1 et de F_2 contrôle des spectres continus à un pic ou à deux pics.
- ii) la compacité correspond à la région à un pic, le caractère diffus à la région à deux pics et l'opposition grave/aigu aux deux possibilités de dominance d'un pic sur l'autre (cf. Fig. 13).

Nous retrouvons ainsi la morphologie catégorisante de type 'point critique' en théorie des transitions de phases que nous avons anticipée en 2.6.2. Rappelons qu'elle ne satisfait pas à la condition de discrétisation.

Il est donc naturel de faire l'hypothèse que la transformation auditive $T : S \rightarrow F$ (cf. 2.6.1) transforme la situation acoustique en la situation limite de la Figure 13 (cf. Fig. 13bis).

Or la situation auditive limite est typiquement catastrophique. Si on substitue aux pics spectraux des minima de potentiel elle correspond à la plus importante des catastrophes élémentaires, à savoir la catastrophe cusp.

La catastrophe cusp déploie une singularité de codimensions 2 de type x^4 c'est-à-dire un minimum dégénéré de potentiel (défini sur un espace interne $M = \mathbb{R}$ de coordonnée x) obtenu par collapse de deux minima et d'un maximum non dégénéré. Le théorème de classification évoqué en 3.6 montre que son déploiement universel est donné par (1) $f_w = x^4 + ux^2 + vx$ les coordonnées u et v parcourant un voisinage W de l'origine de \mathbb{R}^2 (espace de contrôle externe).

Il est facile de vérifier que l'ensemble catastrophique K de W défini comme l'ensemble des $w \in W$ pour lesquels f_w admet (d'après le théorème de Morse) des

FIGURE 13 : Le triangle fondamental /a/ — $\begin{matrix} /u/ \\ /i/ \end{matrix}$ comme catégorisation de type 'point critique'

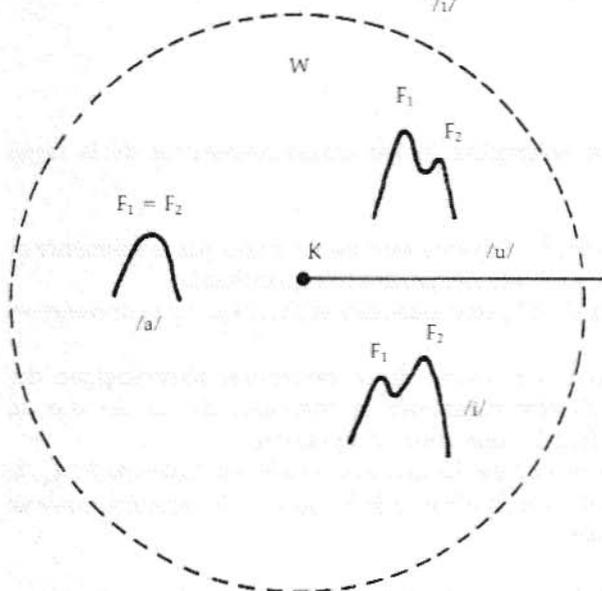
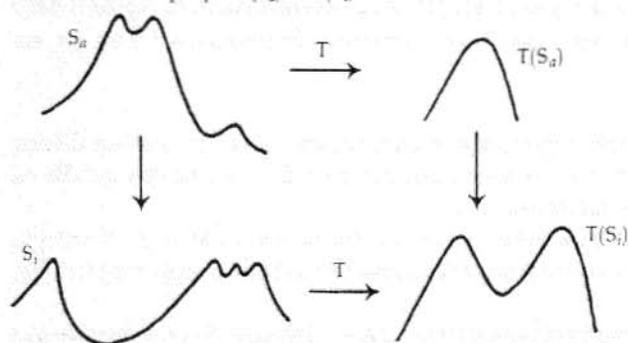


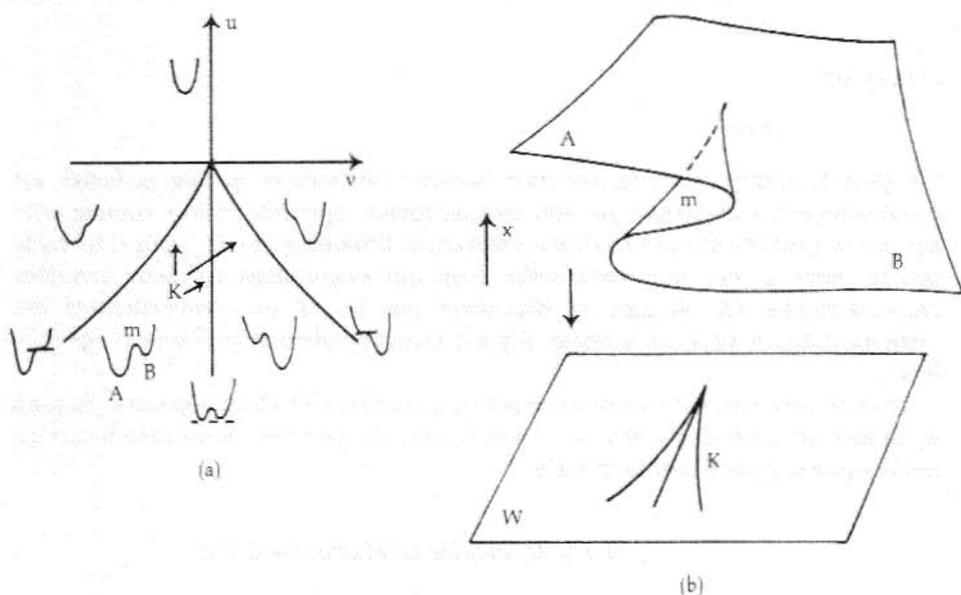
FIGURE 13bis : La transformation T retient l'essentiel des propriétés qualitatives des spectres S_a et S_i de /a/ et de /i/. Elle transforme l'opposition relative 'F₁ et F₂ voisins/F₁ et F₂ séparés' en une opposition absolue 'un maximum/deux maxima'. Par là-même elle transforme la transition continue $S_a \rightarrow S_i$ en une transition catastrophique $T(S_a) \rightarrow T(S_i)$.



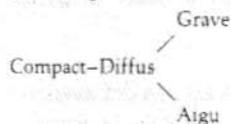
points critiques dégénérés ou des valeurs critiques multiples est celui représenté à la Figure 14(a). Et si l'on porte sur un troisième axe (coordonnée x) les points critiques de f_w on obtient la célèbre surface froncée représentée à la Figure 14(b).

On peut donc faire l'hypothèse qu'après la transformation auditive l'interface /i///u/ est une interface symétrique induite par une catastrophe de conflit alors que les interfaces /a///u/ et /a///i/ sont des interfaces asymétriques induites par des catastrophes de bifurcation. Cette catégorisation de type 'point critique' s'oppose radicalement à celles de type point triple décomposant W en trois domaines par l'intermédiaire de trois strates de conflit. Dans la mesure où ces deux catégorisations

FIGURE 14 : La catastrophe cusp et la surface fronce.



sont les plus typiques (les plus archétypales) pouvant apparaître stablement sur un espace externe de dimension 2, il serait essentiel de pouvoir les tester expérimentalement. Mais, quoi qu'il en soit, la méthodologie catastrophique permet d'affirmer *a priori* que si la description

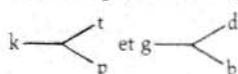


est valide c'est nécessairement une catastrophe cusp qui organise le triangle des voyelles cardinales.

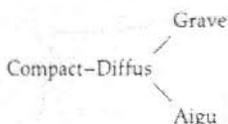
À partir de là, il faudrait pouvoir rendre compte des complexifications de cette catégorisation universelle de base K_0 et retrouver la classification des systèmes vocaliques (cf. Crothers 1978). Pour cela, il faut tenir compte du troisième formant F_3 et donc introduire la catastrophe élémentaire papillon classifiant les potentiels à trois minima et deux maxima. Cette catastrophe de centre organisateur x^6 est de codimension 4. Sa géométrie est déjà d'une complexité notable et suffit à notre avis pour géométriser les hiérarchies vocaliques des diverses langues (cf. Petitot 1982).

4.2 La classification, la hiérarchie et la stratification des occlusives.

On sait que les triangles des occlusives



sont eux aussi traditionnellement décrits par la hiérarchie



On peut le légitimer en faisant avec Stevens l'hypothèse qu'une occlusive est essentiellement caractérisée par son attaque (onset) spectrale conçue comme configuration spectrale d'ordre supérieur (Stevens et Blumstein, 1978). Mais si tel est le cas ce serait encore une catastrophe *cusp* qui organiserait les deux triangles consonantiques. Or Stevens et Blumstein ont trouvé expérimentalement des catégorisations de type *point triple*. Il y a là une contradiction qu'il serait urgent de lever.

Quoi qu'il en soit, montrons ce que peut apporter la méthodologie catastrophique à un modèle précis de la classification des occlusives. Nous avons choisi comme test un modèle proposé par Massaro et Oden.

4.2.1 Le modèle de Massaro – Oden

Dans les deux articles que nous prenons pour fil directeur (Oden et Massaro 1978, Massaro et Oden 1980), les auteurs proposent et testent un modèle d'intégration des indices acoustiques en un percept unifié. Les deux indices qu'ils étudient sont ceux du voisement et de la place d'articulation pour les occlusives. Ils se situent donc dans l'espace de contrôle (v, a) de coordonnées v pour le VOT et a pour la place d'articulation. Leurs hypothèses sont les suivantes :

- i) Les traits acoustiques sont détectés et perçus indépendamment les uns des autres.
- ii) Chaque trait est évalué par une mesure de son degré de présence dans le signal.
- iii) Chaque phonème est défini par un prototype stocké prédictivement dans la mémoire à long terme (intégration de l'information acoustique).
- iv) Chaque son est identifié sur la base de l'évaluation de ses traits par comparaison avec les prototypes en compétition.

Soit s un stimulus phonétique, en l'occurrence une occlusive, et soient v et a les deux traits considérés. Selon Massaro et Oden, les traits sont des propriétés de s qui ne sont pas binaires mais possédées par s avec un certain degré et qui sont décrites en conséquence par des prédicats flous. Autrement dit, si l'on considère les oppositions voisé/~voisé et labial/~labial (alvéolaire) on leur associera les prédicats flous V et L et l'on aura donc suivant les règles du calcul des prédicats flous :

- i) $V(s), L(s) \in [0, 1]$,
- ii) $\sim V(s) = 1 - V(s)$, etc.
- iii) $V \wedge L(s) = V(s) \cdot L(s)$, etc.

Dans une première élaboration de leur modèle, Massaro et Oden ont simplement identifié les prototypes des occlusives aux prédicats flous traduisant leur caractérisation booléenne en termes de traits distinctifs (\pm , \pm). Ils ont donc utilisé les correspondances :

$$/b/ \Leftrightarrow B = V \cdot L$$

$$/d/ \Leftrightarrow D = V \cdot (1 - L)$$

$$/p/ \Leftrightarrow P = (1 - V) \cdot L$$

$$/t/ \Leftrightarrow T = (1 - V) \cdot (1 - L)$$

Si $|\alpha|$ est un des prototypes considérés et si l'on définit la probabilité que le stimulus s soit identifié à $|\alpha|$ à la probabilité conditionnelle

$$p(|\alpha|/s) = \frac{A(s)}{B(s) + D(s) + P(s) + T(s)}$$

où $A(s)$ est le prédicat associé à $|\alpha|$, on obtient tout simplement $p(|\alpha|/s) = A(s)$ puisque $VL + V(1 - L) + (1 - V)L + (1 - V)(1 - L) = 1$.

Il est facile de calculer ce que prédit ce modèle élémentaire. Plaçons-nous dans le plan de contrôle $W = (v, a)$, $v, a \in [0, 1]$. Nous pouvons identifier l'évaluation de s à sa position dans W puisque $V(s) = v(s)$ et $L(s) = a(s)$. Étant donné un point s de W nous calculons les valeurs en s des quatre 'prédicats' (i.e. des quatre polynômes du second degré) va , $v(1 - a)$, $(1 - v)a$ et $(1 - v)(1 - a)$ et nous posons que s appartient à la catégorie donnant la plus grande valeur. L'interface K catégorisant W est donc constituée des points (singuliers) de W où il y a compétition entre deux prédicats et elle est définie par les équations :

$$va = v(1 - a) \quad \text{i.e. } a = \frac{1}{2} \quad \text{ou } v = 0$$

$$va = (1 - v)a \quad \text{i.e. } v = \frac{1}{2} \quad \text{ou } a = 0$$

$$va = (1 - v)(1 - a) \quad \text{i.e. } v + a = 1$$

$$v(1 - a) = (1 - v)a \quad \text{i.e. } v = a$$

$$v(1 - a) = (1 - v)(1 - a) \quad \text{i.e. } v = \frac{1}{2} \quad \text{ou } a = 1$$

$$(1 - v)a = (1 - v)(1 - a) \quad \text{i.e. } a = \frac{1}{2} \quad \text{ou } v = 1$$

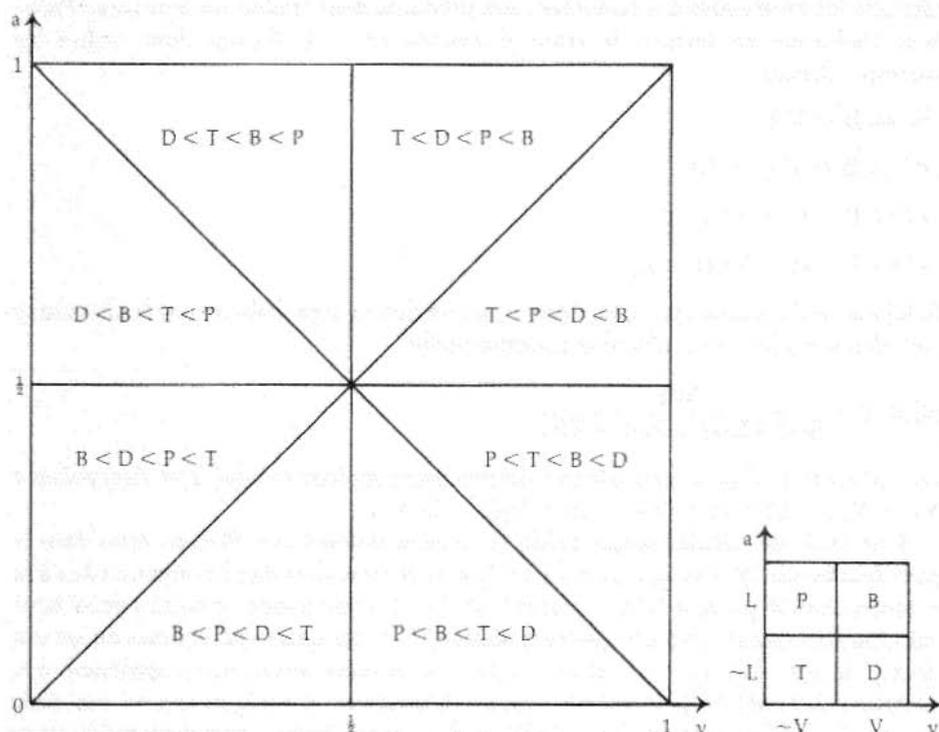
Les 4 droites $v = \frac{1}{2}$, $a = \frac{1}{2}$, $v + a = 1$ et $v = a$ partitionnent N en 8 domaines et il est trivial de calculer l'ordre de B , D , P et T dans chacun de ces domaines à partir des inégalités :

$$v \geq \frac{1}{2} \Leftrightarrow B \geq P \text{ et } D \geq T$$

$$a \geq \frac{1}{2} \Leftrightarrow B \geq D \text{ et } P \geq T$$

$$v + a \geq 1 \Leftrightarrow B \geq T$$

$$v \geq a \Leftrightarrow D \geq P \quad (\text{cf. Fig. 15})$$

FIGURE 15 : Catégorisation de W induite par le modèle élémentaire de Massaro et Oden.

D'autre part, le modèle prédit que les courbes de probabilité $p(|\alpha|/s)$ en fonction, par exemple, de a pour v fixé sont des droites. Or les résultats des tests d'identification ne vérifient pas cette prédiction (cf. Fig. 16).

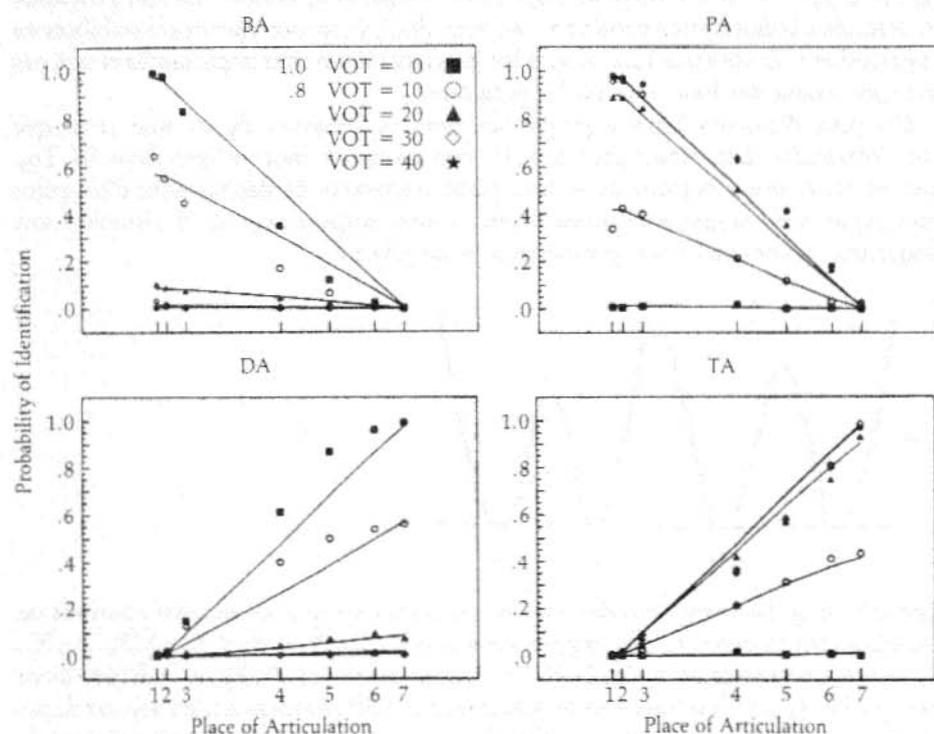
Ce désaccord a conduit Massaro et Oden à complexifier leur modèle en posant que les prédicats flous définissant les prototypes des occlusives n'étaient pas simplement les traductions de leur description en termes de traits distinctifs mais étaient des fonctions générales des traits, par exemple de la forme $A = V^p L^q$, ou $A = (1 - V)^p L^q$, ou $A = V^p + L^q$, etc. Le problème devient alors de trouver les fonctions qui concordent le mieux avec les résultats expérimentaux (cf. Fig. 17).

De telles complexification du modèle conduisent à une transformation de l'interface K (cf. Fig. 18) et c'est sur ce dernier point que nous voudrions faire quelques remarques. Ce modèle soulève en effet des problèmes catastrophiques qu'il serait intéressant de développer. Son idée générale est la suivante. On considère un espace de contrôle W dont les coordonnées $w = (u_1, \dots, u_r)$ (ici (v, a)) sont des valeurs d'indices acoustiques. On se donne ensuite n prototypes définis par n fonctions $P_1(w), \dots, P_n(w)$ (ici B, D, P et T). Ces prototypes définissent un champ

$$\sigma : W \rightarrow \mathbb{R}^n$$

$$w \rightarrow (P_1, \dots, P_n).$$

FIGURE 16 : D'après Oden et Massaro (1978 : 181). Identification probabilities and predictions of the simple fuzzy logical model. (Each panel presents the data for a given response alternative. Note that the spacing along the abscissa is proportional to the spacing of the subjective place values. VOT = voice onset time.)



Dans \mathbb{R}^n les égalités $P_i = P_j$ ($i \neq j$) déterminent une stratification K_0 qui est l'image inverse par l'application

$$\gamma_n : \mathbb{R}^n \rightarrow \mathbb{R}^N$$

$$(P_i) \rightarrow (P_i - P_j)_{i > j}$$

de la stratification

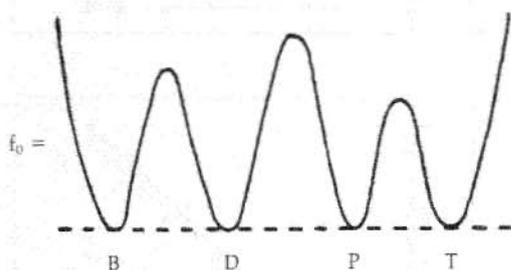
$$K_* \text{ de } \mathbb{R}^N \left(N = \frac{n(n-1)}{2} \right)$$

définie par les hyperplans de coordonnées. La catégorisation de W par un système K d'interfaces n'est alors rien d'autre que l'image inverse de K_0 par σ i.e. de K par $\gamma_n \circ \sigma$.

Le modèle 'simple' dérive d'un champ $\sigma_0 : W \rightarrow \mathbb{R}^n$ qui n'est pas transversal sur la stratification K_0 . Cette situation structurellement instable est prise pour 'centre organisateur' de modèles 'complexes' qui en sont des stabilisés et où σ devient

transversal sur K_0 et évite donc les strates de K_0 de codimension >1 . Cette remarque est épistémologiquement et méthodologiquement importante car elle montre qu'il est vain de chercher une expression exacte des fonctions P_i caractérisant prédictivement les prototypes. La bonne méthodologie consiste à partir de la situation non générique σ_0 associée à la description booléenne (au sens 'flou') des prototypes et à la stabiliser en la perturbant. Encore une fois, ce sont les interfaces K qui sont importantes et non pas la forme exacte des fonctions qui les induisent.

On peut d'ailleurs faire à propos de cette stabilisation de σ_0 une remarque supplémentaire. L'interface générique K obtenue dans le modèle 'complexe' (cf. Fig. 18) est un exemple typique de section plane transverse du déploiement d'un point quadruple. Cela suggère de prendre pour centre organisateur de la situation une singularité point quadruple, par exemple la singularité :



Une telle singularité est de codimension 3 et il est clair que ses stabilisés partiels ou complets sont exactement classés par les situations $X_1 \leq X_2 \leq X_3 \leq X_4$ où (X_1, X_2, X_3, X_4) est une permutation de (B, D, P, T) . Autrement dit, si (\wedge, Γ) est le déploiement universel de f_0 et si l'on considère le champ $\psi : \wedge \rightarrow \mathbb{R}^4$ associant à f_x les valeurs de ses minima, l'ensemble catastrophique Γ est l'image inverse de K par l'application $\gamma_4 \circ \psi$.

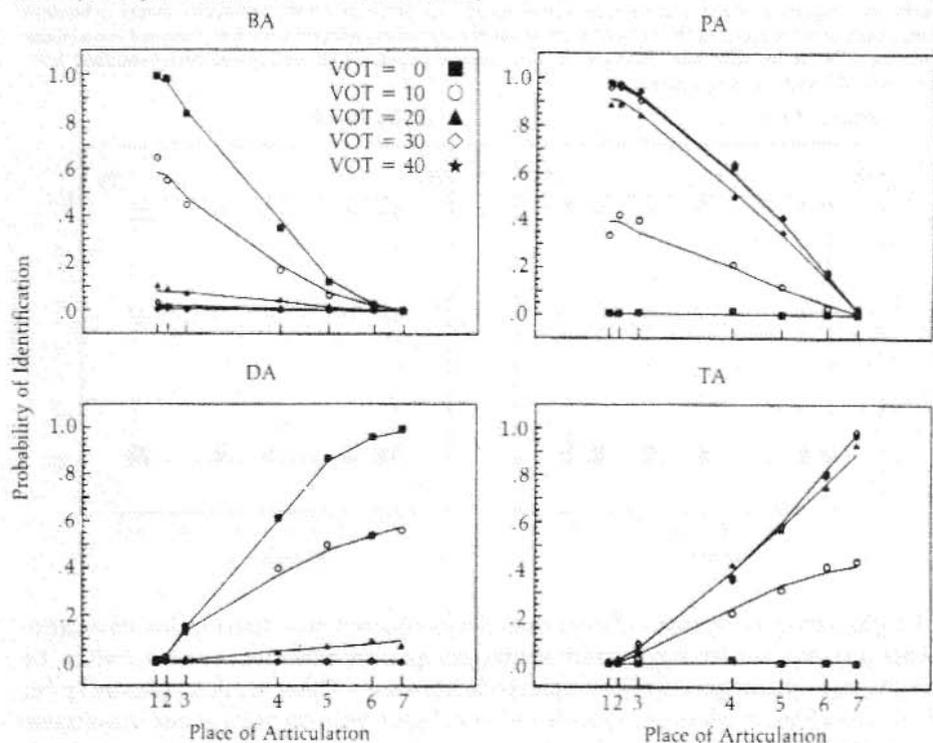
Il serait donc intéressant de savoir si le champ $\sigma : W \rightarrow \mathbb{R}^4$ cherché par Massaro et Oden peut se factoriser à travers ψ . Une telle factorisation est loin d'être évidente dans la mesure où elle présuppose que l'on sache associer à chaque stimulus s de W un potentiel f_s à 4 minima dont les valeurs seraient les évaluations $B(s)$, $D(s)$, $P(s)$ et $T(s)$. Or une telle fonction globale (que l'on peut toujours introduire abstraitement) serait définie sur un espace interne ininterprétable.

Acceptons toutefois l'idée d'une factorisation. On peut alors beaucoup mieux visualiser la stabilisation de σ_0 . En effet σ_0 (décrit à la figure 15) correspond à une section plane non générique (puisque passant par le centre organisateur qui est de codimension $3 > 2$) du déploiement universel (\wedge, Γ) de f_0 . Cette section est représentée à la figure 19.

La non généricité de σ_0 est évidente une fois ainsi exprimée. W étant en effet de dimension 2, il ne peut présenter stablement que des catastrophes de conflit de complexité inférieure à celle des points triples. Tel n'est pas le cas pour σ_0 . Dans (W, K) les axes de coordonnées v et a sont des lignes de doubles conflits, alors que les doubles conflits étant de codimension 2 ils ne peuvent apparaître stablement en dimension 2 que sous forme de points isolés.

Lorsque σ_0 se stabilise, W devient une section générique de (\wedge, Γ) , le point

FIGURE 17 : D'après Oden et Massaro (1978 : 183). Identification probabilities and predictions of the complex fuzzy logical model. (Note the the spacing along the abscissa is proportional to the spacing of the subjective place values. VOT = voice onset time.)

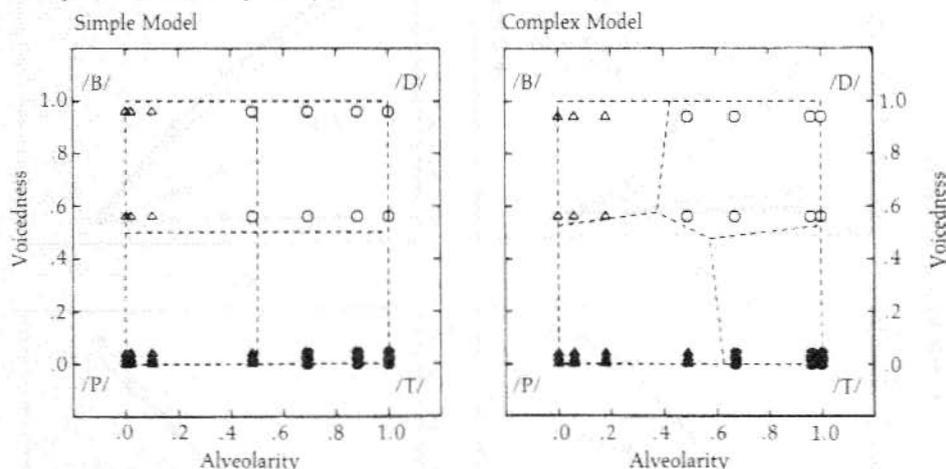


quadruple 'explose' en 4 points triples et un point de double conflit et les axes se dédoublent conformément aux deux possibilités représentées à la figure 20. En appliquant la convention de Maxwell, l'on obtient l'ensemble catastrophique du modèle 'complexe' de Massaro-Oden. On trouvera à la figure 21 la classification des types différenciables opérée par les deux types de sections génériques de la figure 20.

4.2.2 Stratification des interfaces et hiérarchie consonnantique.

Indépendamment du contenu computationnel du modèle censé en rendre compte, les expériences de Massaro et Oden montrent que, dans le plan (v, a) , le système d'interfaces faisant des occlusives /b/, /d/, /p/ et /t/ des valeurs positionnelles est une section générique d'un point quadruple c'est-à-dire un archétype de décomposition d'un plan en 4 régions par des strates de Maxwell. Pour être complet il faudrait prolonger les expériences à l'ensemble des 6 occlusives de base /b - d - g/ et /p - t - k/. On peut penser qu'on obtiendrait alors un ensemble catastrophique recollant deux modèles, celui de /b, d, p, t/ d'une part et celui de /d, t, g, k/ d'autre part (cf. Fig. 22), ce qui montrerait d'ailleurs à quel point les expériences de Lisker et Abramson étaient, ainsi que nous l'avons noté, insuffisantes.

FIGURE 18 : Transformation de l'interface K catégorisant W par complexification des prédicats flous définissant les prototypes des occlusives. D'après Oden et Massaro (1978 : 185). Psychological parameter spaces for the simple and complex fuzzy logical models. (The dashed lines partition the parameter spaces into regions in which a given response alternative is predicted by the respective model to be most likely. Each point represents the subjective position of a stimulus and specifies what phoneme it was most often identified to be, with open triangles, closed triangles, open circles, and closed circles standing for /b/, /p/, /d/, and /t/, respectively.)



La géométrie du système d'interfaces des occlusives peut fournir des renseignements précieux sur les rapports hiérarchiques qu'elles entretiennent entre elles. Le fait que, dans la catégorisation du modèle de Massaro - Oden, les domaines de /p/ et /d/ soient adjacents alors que ceux de /b/ et /t/ sont disjoints indique que le contraste est plus fort entre /b/ et /t/ qu'entre /p/ et /d/.

Aussi naïve soit-elle, cette remarque permettrait peut-être d'interpréter en termes de stratifications phonétiques (en donc en termes de phénoménologie de la substance) les universaux hiérarchiques phonologiques dégagés dans le système des occlusives (ou des fricatives et des voyelles). Ainsi pourrait se développer une théorie catastrophiste des échelles de marquage.

Thomas Gamkrelidze a par exemple montré que si le voisement est pertinent pour les occlusives d'un système phonologique, alors (1) le /b/ voisé + labial domine le /g/ voisé + vélaire et (2) le /k/ non voisé + vélaire domine le /p/ non voisé + labial (Gamkrelidze 1972 et 1978). Ces rapports de dominance manifestent un universel phonologique selon lequel, lorsqu'un système d'occlusives est incomplet, ce sont toujours les occlusives recessives /p/ et /g/ qui font défaut. On peut penser qu'un tel universel n'est que la manifestation phonologique de la géométrie de l'ensemble catastrophiste de la figure 22. Si une telle hypothèse peut être confirmée, la perception catégorielle ne se bornerait pas à rendre compatibles localement les niveaux audio-acoustiques et phonologiques de la phonétique en ce qui concerne la notion de trait distinctif. Elle les rendrait compatibles globalement en ce qui concerne la notion de système phonologique. Via le paradigme catastrophiste, un système phonologique apparaîtrait alors comme un système global d'inter-

FIGURE 19 : Le modèle 'simple' de la figure 15 comme section non générique du déploiement universel d'un point quadruple f_0 .

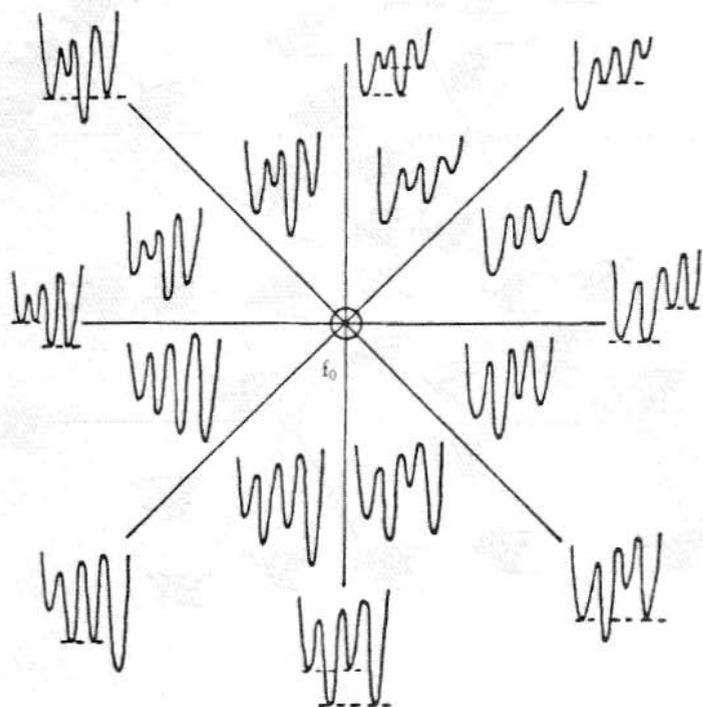


FIGURE 20 : Les deux stabilisations de la section non générique de la figure 19. Les strates de conflit associées à la convention de Maxwell sont en traits pleins. Les cercles représentent les points triples et les carrés les points de double conflit.

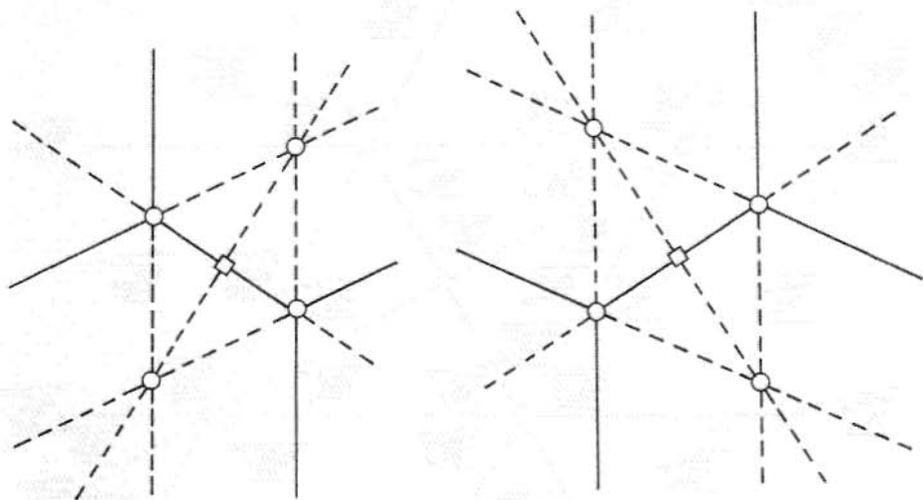


FIGURE 21(b)

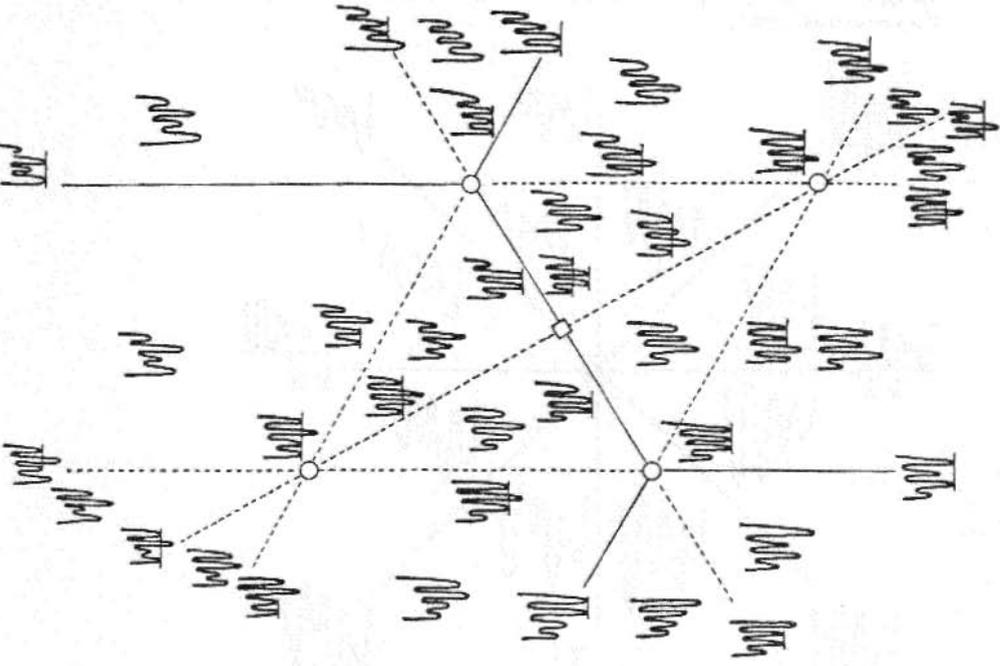


FIGURE 21(a)

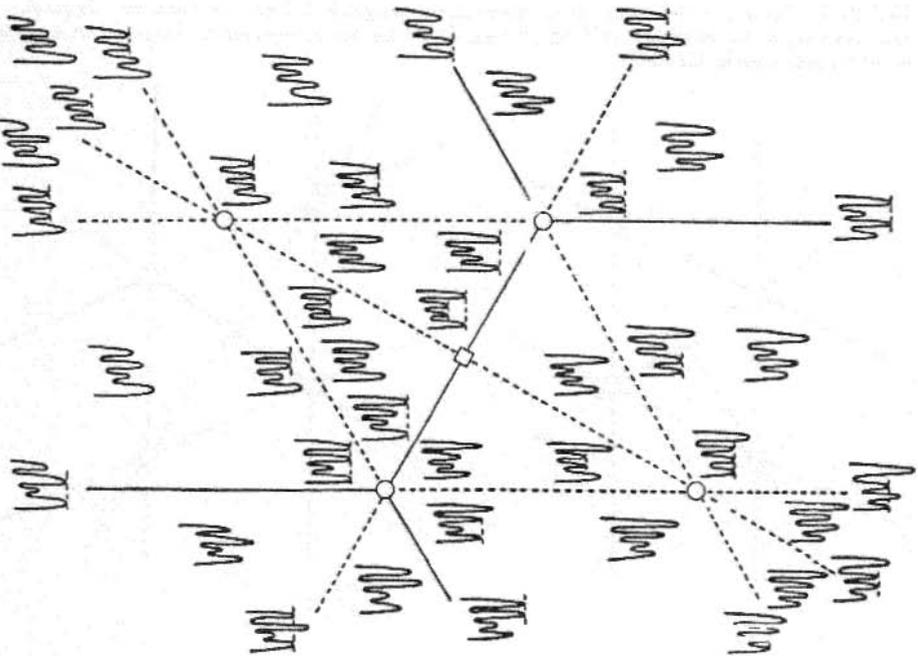
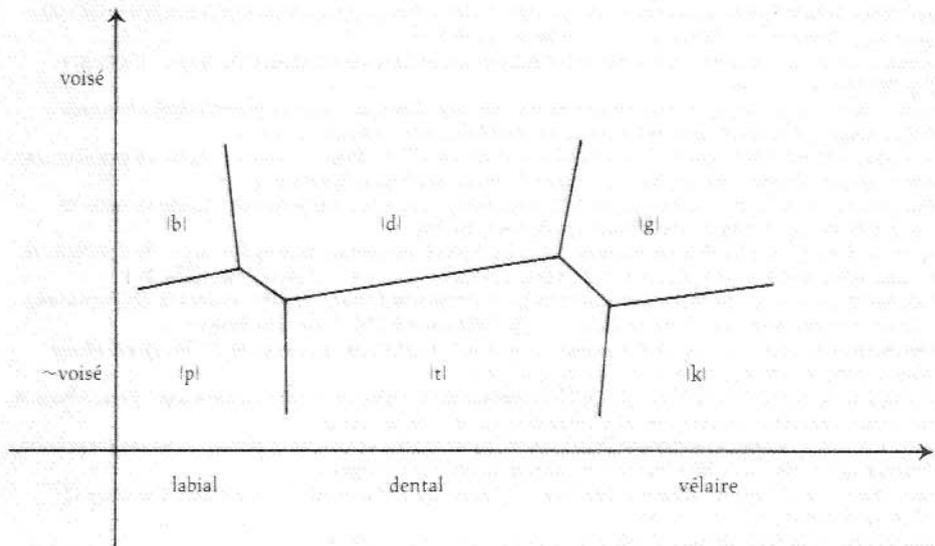


FIGURE 22 : Proposition pour les interfaces des occlusives.



faces phonétiques dont la stratification exprimerait les relations hiérarchiques de marquage.

NOTES

- 1 Sur le conflit inné/acquis en psycholinguistique, cf. le débat Chomsky-Piaget (Piattelli-Palmarini, 1979)
- 2 Ces théories motrices admettent corrélativement que l'articulation est phonologiquement finalisée à travers l'internalisation d'un 'espace' de cibles phonémiques (Mac Neilage, 1979).
- 3 Autrement dit, le substrat psychophysique est une cause matérielle de la forme relationnelle de l'expression sans en être pour autant la cause efficiente et/ou formelle.
- 4 Pour des précisions mathématiques, cf. par exemple Thom 1972, 1980, Zeeman 1977, Golubitsky et Guillemin 1973, Chenciner 1973, 1980 et Petitot 1982.

RÉFÉRENCES

- ABBS, J.H. & H.M. SUSSMAN (1971) 'Neurophysiological feature detectors and speech perception: a discussion of theoretical implications,' *Journal of Speech and Hearing Research*, 14, 1: 23-36.
- BRYANT, J.S. (1978) 'Feature Detection Process in Speech Perception,' *Journal of Experimental Psychology*, (Human Perception and Performance), 4, 4: 610-620.
- CHENCINER, A. (1973) 'Travaux de Thom et Mather sur la stabilité topologique,' *Séminaire Bourbaki*: 424.
- (1980) 'Singularité des fonctions différentiables,' *Encyclopedia Universalis*.
- CROTHERS, J. (1978) 'Typology and Universal of Vowel Systems,' *Universals of Human Language*, (J.H. Greenberg, ed.) Vol. 2, Stanford University Press: Stanford: 93-152.
- CUTTING, J.E. & B.S. RÖSNER (1974) 'Categories and boundaries in speech and music,' *Perception and Psychophysics*, 16, 3: 564-570.
- DIEHL, R.L. ET AL. (1978) 'Contrast Effects on Stop Consonant Identification,' *Journal of Experimental Psychology*, (Human Perception and Performance), 4, 4: 599-609.

- EIMAS, P.D. (1974) 'Auditory and linguistic processing of cues for place of articulation by infants,' *Perception and Psychophysics*, 16, 3: 513-521.
- (1983) 'Infant Speech Perception: Issues and Models,' *Perspectives on mental Representation*, (M. Garrett, J. Mehler, E. Walker, eds.) L. Erlbaum: Hillsdale, N.J.
- EIMAS, P.D. & J.D. CORBITT (1973) 'Selective Adaptation of Linguistic Feature Detectors,' *Cognitive Psychology*, 4: 99-109.
- GAMKRELIDZE, I.V. (1972) 'Hierarchical relations among phonemic units as phonological universals,' *Proceedings of the ninth international Congress of Phonetic Sciences*, 2: 9-15.
- (1978) 'On the Correlation of Stops and Fricatives in a Phonological System,' *Universals of Human Language*, (J. Greenberg ed.), Vol. 2, Stanford University Press: Stanford: 9-46.
- GOLUBITSKY, M. & V. GUILLEMIN (1973) *Stable Mappings and Their Singularities*, Graduate texts in mathematics, 14, Springer: New York, Heidelberg, Berlin.
- JUSCZYK, P.W. (1983) 'Is there a Phonetic Basis for Speech Perception During Infancy?' *Perspectives on mental representation* (M. Garrett, J. Mehler, E. Walker, eds.), L. Erlbaum: Hillsdale, N.J.
- LADEFOGED, P. (1972) 'Phonetic prerequisites for a distinctive feature theory,' *Papers in Linguistic and Phonetics to the memory of Pierre Delattre*, (A. Valdman ed.) Mouton: The Hague.
- LIBERMAN, A.M., F.S. COOPER, D.P. SHANKWEILER, & M. STUDDERT-KENNEDY (1967) 'Perception of the speech code,' *Psychological Review*, 74, 6: 431-461.
- LINDBLOM, B. (1972) 'Some Phonetic null hypotheses for a biological theory of language,' *Proceedings of the ninth international Congress of Phonetic Sciences*, Vol. 2: 33-40.
- LISKER, L. & A.S. ABRAMSON (1970) 'Discriminability along the voicing continuum: cross language texts,' *Proceedings of the sixth International Congress of Phonetic Sciences*.
- MAC NEILAGE, P.F. (1979) 'Speech production,' *Proceedings of the Ninth International Congress of Phonetic Sciences*, Vol. 1: 11-39.
- MALMBERG, B. (1974) *Manuel de phonétique générale*, Picard: Paris.
- MASSARO, D.W. (1972) 'Issues in speech perception,' *Proceedings of the ninth International Congress of Phonetic Sciences*, Vol. 2: 474-481.
- MASSARO, D.W. & G.C. ODEN (1980) 'Evaluation and integration of acoustic features in speech perception,' *Jour. Acoust. Soc. Am.*, 67, 3: 996-1013.
- MEHLER, J. & J. BERTONCINI (à paraître) 'Infants' Perception of Speech and other Acoustic Stimuli,' *Infant behaviour and Development*, 2.
- MILLER, G.A. & P.H.N. JOHNSON-LAIRD (1976) *Language and Perception*, Harvard University Press: Cambridge.
- MILLER, J.D. ET AL. (1976) 'Discrimination and labelling of noise-buzz sequences with varying noise-lead times: an example of categorical perception,' *Jour. Acoust. Soc. Am.*, 60: 410-417.
- ODEN, G.C. & D.W. MASSARO (1978) 'Integration of Featural Information in Speech Perception,' *Psychological Review*, 85, 3: 172-191.
- PASTORE, R.E. ET AL. (1977) 'Common-Factor Model of Categorical Perception,' *Journal of Experimental Psychology*, (Human Perception and Performance), 3, 4: 686-696.
- PETITOT, J. (1982) *Pour un Schématisme de la Structure: de quelques implications sémiotiques de la théorie des catastrophes.*, Thèse, Ecole des Hautes Etudes en Sciences Sociales: Paris.
- PIATTELLI-PALMARINI, M. (1979) *Théories du langage. Théories de l'apprentissage. Le débat entre Jean Piaget et Noam Chomsky.* Seuil: Paris
- PISONI, D.B. (1979) 'On the Perception of Speech Sounds as Biologically Significant Signals,' *Brain Behav. Evol.*, 16: 330-350.
- REPP, B.H. ET AL. (1978) 'Perceptual Integration of Acoustic Cues for Stop, Fricative, and Affricate Manner,' *Journal of Experimental Psychology*, (Human Perception and Performance), 4, 4: 621-637.
- SIEGEL, J.A. & W. SIEGEL (1977) 'Categorical perception of tonal intervals: Musicians can't tell sharp from flat,' *Perception and Psychophysics*, 21, 5: 399-407.
- STEVENS, K.N. (1972) 'The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data,' *Human Communication, a Unified View*, (P.B. Denes, E.E. David Jr. eds.).
- STEVENS, K.N., & S.E. BLUMSTEIN (1978) 'Invariant cues for place of articulation in stop consonants,' *Jour. Acoust. Soc. Am.*, 64, 5: 1358-1368.
- STUDDERT-KENNEDY, M., ET AL. (1970) 'Motor theory of speech perception: a reply to Lane's critical review,' *Psychological Review*, 77-3: 234-249.
- THOM, R. (1972) *Stabilité Structurale et Morphogénèse*, Benjamin: New-York; Ediscience: Paris.
- (1980) *Modèles Mathématiques de la Morphogénèse*, (2ème ed.), Christian Bourgois: Paris.
- ZEEMAN, C. (1977) *Catastrophe Theory*, Addison-Wesley: Reading, MA.