

## Publié dans Faits de Langues (1999)

Vaissière, J., (1999), "[Utilisation de la prosodie dans les systèmes automatiques](#) : un problème d'intégration des différentes composantes.", [Faits de Langues](#), Oral-écrit: Formes et théories, Ophys, 9-16

### Utilisation de la prosodie dans les systèmes automatiques : un problème d'intégration des différentes composantes.

Jacqueline Vaissière 


#### ABSTRACT

It has been tried in the past to apply prosodic knowledge in automatic speech synthesis and recognition. The passage from read speech to spontaneous speech requires the understanding of the interactions among the different components, 1) syntactic, 2) semantico-pragmatic, 3) rhythmic, 4) lexico-morphological, 5) attitudinal and emotional, 6) stylistic, 7) individual, 8) physiological, and 9) rate-related components. The use of statistics on large data base is difficult because of the large number of components and the lack of a satisfactory method to annotate prosodically the speech material.

#### RESUME

On a essayé d'utiliser dans le passé les connaissances prosodiques dans l'apprentissage et la reconnaissance et en synthèse automatiques de la parole. Le passage de la parole lu au spontané nécessite une compréhension de l'interaction entre les différentes composantes : 1) syntactique, 2) sémantico-pragmatique, 3) rythmique, 4) attitudinelle et émotionnelle, 5) stylistique, 7) individuelle, 8) physiologiques, 9) débit. Les méthodes statistiques sont difficiles d'utilisation à cause du nombre élevé de composantes, et du manque d'une méthode d'annotation prosodique des bases de données.

---

 Université de la Sorbonne Nouvelle, UPRESA 7018, jvaiss@msh-paris.fr

## MOTS CLES :

Prosodie , syntaxe, énonciation, applications technologiques

## 1. INTRODUCTION : UN ESSAI DE DEFINITION DE LA PROSODIE

La compréhension du système prosodique inclut, entre autres, celle de sa production, sa perception, et ses aspects phonologique et fonctionnel.

Sur le plan de la production, la prosodie concerne l'ensemble des *variations acoustiques* dans la réalisation des phonèmes qui ne dépendent pas de leur identité et de celle de leur entourage phonétique immédiat. Les sept types de variations les plus connus sont : des coefficients d'allongements (ou raccourcissements) des unités phoniques (par rapport, par exemple, à leur longueur moyenne observée dans une grande base de données, Campbell, 1992), des fluctuations de la fréquence du fondamental (Fo), les pauses , des variations d'intensité et de pente spectrale, des modifications de timbre (allophones prosodiques), des ajouts (« consonnes » de liaison, insertion d'un coup de glotte, « voyelle » ou « consonne » épenthétiques) ou des élisions de sons.

Sur le plan de la perception, il n'est pas prouvé que toutes les variations acoustico-prosodiques sont porteuses d'information. Un mécanisme de filtrage pas encore entièrement élucidé permet à l'auditeur de dégager des variations les corrélats d'un certain nombre de « traits prosodiques », qui donnent lieu à des impressions perceptives soit discrètes (accent d'insistance, focalisation, modalité, jointure, etc.), soit continues (différents degrés d'accent et de jointure, nuances d'attitude, etc.)

L'approche dite phonologique se situe dans la lignée de la grammaire générative, mais l'accent est mis désormais sur la représentation (multilinéaire) des unités, plutôt que sur l'élaboration de règles et de grammaires prosodiques. La prosodie se décrit comme les structures de niveaux supérieurs aux phonèmes délimitées par les variations prosodiques : la rime, la syllabe, le pied, le groupe accentuel (« stress group »), les mot, syntagme et énoncé prosodiques, etc. (les termes varient selon les descriptions).

L'approche fonctionnelle de la prosodie se situe dans la continuation des travaux de l'école de Prague. La langue est analysée en priorité comme un système d'unités fonctionnelles : expression de la modalité, fonction de segmentation du continuum et de hiérarchisation, et fonction expressive.

Les variations généralement intégrées dans les systèmes automatiques se réfèrent aux modèles de production et se limitent aux variations de Fo, de durée, et les pauses.

## 2. LES DIVERSES COMPOSANTES

La complexité de la parole a obligé les phonéticiens et phonologues à créer des paradigmes spécifiques, parfois artificiels, pour dégager les composantes qui aboutissent à des variations acoustico-prosodiques. Neuf composantes (au moins) ont été dégagées. Faute de place, il nous est impossible de citer les travaux les plus importants dans chacun des domaines (Nous référons à Vaissière, 1989, pour une bibliographie partielle).

### *2.1. la composante syntaxique*

La structure syntaxique impose des contraintes sur la structuration prosodique. Par exemple, les mots grammaticaux sont généralement marqués par un passage dans le registre bas du locuteur (Vaissière, 1971) et un passage dans le registre haut signale la présence sous-jacente d'un mot prosodique (composé par un ou deux mots lexicaux unis par sens); la distance acoustique entre les mots prosodiques en séquence tend à refléter leur dépendance syntaxique, etc. (cf. Delattre, 1961, pour la notion de continuation majeure et mineure). L'étude des rapports prosodie/syntaxe a été favorisée par l'approche générative. Le meilleur paradigme pour mettre en valeur ces rapports est l'étude de textes lus de façon « neutre », voire de phrases isolées. Malgré la non-congruence parfaite entre prosodie et syntaxe (reconnue très tôt, dès 1971, pour le français), le succès de cette approche s'est concrétisé par la génération de la prosodie en synthèse de parole lue. Le lecteur pourra consulter le numéro spécial de la revue ATALA (1997), sur « Prosodie et Syntaxe », consacrée à cette composante.

### *2.2. la composante énonciative et les règles sémantico-pragmatiques*

Avant les années 70, les rapports entre prosodie et la structure informationnelle dans les énoncés spontanés ont fait l'objet de travaux qui n'ont rien perdu de leur actualité. La frontière acoustique majeure marque l'articulation thème/rhème, plutôt que la structure syntaxique au sens stricte. Une focalisation acoustique perturbe très fortement les profils mélodiques et rythmiques. L'approche énonciative favorise le point de vue des utilisateurs dans l'acte global de communication, l'effet des contraintes et interactions sociales, l'étude des marques de la progression du discours, etc. C'est très certainement la plus importante et la plus difficile des composantes à étudier, car la parole est majoritairement spontanée (et non lue) et la structure informationnelle prime sur la structure syntaxique dans le spontané.

### *2.3. la composante métrique (rythmique)*

Le profil prosodique peut être plus ou moins fortement perturbé par des tendances rythmiques : tendance à l'alternance locale entre syllabes fortes et syllabes, qui conduit à certains « déplacements d'accents » en anglais et en français (« thir'teen + 'men » > « 'thirteen 'men » ; « café noir/café au lait ») (cf. la grille métrique de Liberman et al, 1977 ; Duez, 1985 pour le français); tendance moins locale à répéter ou alterner les mêmes patrons de mot prosodique, montants ou descendants, indépendamment des structures syntaxiques et

informationnelles (Vaissière, 1974 et 1975, Martin, 1982), etc. L'interaction des composantes 1 et 3 est magistralement illustrée dans la structure de performance (Monnin et al, 1993).

#### 2.4. la composante lexico-morphologique

C'est cette composante qui semble varier le plus selon les langues (Vaissière, 1983). Par exemple, la composition morphologique du mot régit entièrement les positions des accents primaire et secondaire en allemand ; elle n'influence pas la position des accents secondaires en anglais, et son influence survit sous la forme de variantes possibles en français (« **inter**département**al**, **re**venir/**inter**département**al**, **re**venir) ». En allemand comme en anglais (des langues dites à accent, ou « stress languages »), l'accent d'insistance se traduit acoustiquement par un renforcement de la syllabe portant déjà un accent lexical. En français, les deux syllabes aux extrémités du mot sont « marquées » (cf. la notion d'arc accentuel de Fonagy et les patrons de mot de Vaissière, 1975). Les deux extrémités peuvent être perçus comme proéminents sous l'effet d'un accent d'insistance, tout au moins dans les mots courts (« **Paris**, **Paris**, le **paradis**, le **paradis** », mais pas « \*le **paradis** »). Si la syllabe finale est exclusivement porteuse des "intonèmes" de continuation (Delattre, 1961), et participe donc directement à la fonction démarcative (marquage de la frontière droite et segmentation en thème/rhème, par exemple), le degré d'élévation de Fo au début d'un mot n'est uniquement lié au degré d'importance du mot dans la structure informationnelle : il peut être aussi un indice du degré de séparation acoustique avec le mot qui précède (une cible de Fo plus haut contribue à indiquer un degré d'indépendance plus grand avec qui précède), et a une fonction démarcative (marquage de la frontière gauche) (cf. voir la notion de « boundary languages », ou langues à frontières, Vaissière, 1991).

#### 2.5. la composante attitudinelle et émotionnelle

La réalisation d'un cliché mélodique sur une partie de l'énoncé ou tout l'énoncé pour exprimer une attitude (doute, etc.) détermine le profil prosodique et diminue (ou supprime) l'influence possible des autres composantes (cf. les travaux de Ivan Fonagy, par exemple Fonagy 1983, et la notion de double encodage de la parole). Les efforts de modélisation de cette composante, encore insuffisants, ont démarré (quelques thèses en France soutenues ou en préparation).

#### 2.6. la composante stylistique

A chaque style correspondent des profils mélodiques et temporelles particuliers qui permettent à l'auditeur, en voix filtrée (donc incompréhensible) de l'identifier (lecture de conte de fée, interview, conférence, spontané, etc.). L'influence du style a été particulièrement (bien) étudiée pour le français par Lucci (1979). Par exemple, la parole adressée au public et la lecture favorisent, par exemple, l'apparition d'accents initiaux, et on observe des différences de registres, de

débit, etc. L'ensemble des règles permettant la transformation d'un style à un autre n'est pas encore connu.

### 2.7. les composantes physiologiques

Les mécanismes respiratoires et les contraintes physiologiques en général imposent de fortes restrictions sur la parole et conduisent à un certain nombre de similarités entre les langues (Vaissière, 1983 et 1995). Ces contraintes deviennent primordiales dans la parole très rapide, où la probabilité d'apparition des pauses, devenues exclusivement respiratoires, croit avec le nombre de syllabes produites depuis la dernière prise d'air, et les pauses surviennent alors à des endroits dont ne peuvent rendre compte les autres composantes.

### 2.8. la composante individuelle

En plus des variations de tessiture vocale dues à des différences anatomiques, deux locuteurs du même idiolecte peuvent préférer statistiquement, surtout à l'intérieur d'un syntagme prosodique, un patron mélodique de mot à un autre (Vaissière, 1975) : montant, à pic, plat ou descendant. Un accent dit « régional » (breton, toulousain, marseillais, lyonnais, etc.) accentue les différences individuelles, et s'il a valeur d'affirmation identitaire, il peut être renforcé par le locuteur. Il manque très certainement d'études systématiques sur les variantes prosodiques du français parlé en France (cf. cependant les études sur les accents régionaux de F. Carton), les français considérant souvent la norme comme seule digne d'intérêt.

### 2.9. L'influence du débit

Toutes choses égales par ailleurs, l'accélération du débit de parole aboutit à une diminution du nombre de niveaux marqués, voire à une réorganisation complète de la structuration prosodique (Vaissière, 1989 ; Fougeron et al, 1998). L'étude systématique de cette composante a à peine démarré.

La simultanéité dans la réalisation des différentes composantes introduit une certaine confusion dans l'interprétation des faits prosodiques. Prenons pour exemple une élévation de  $F_0$  à l'initiale de mot, sans augmentation sensible de la durée et de l'intensité. Pour Léon (1971), ce type de « joncture » démarcative externe a un rôle secondaire sur le plan référentiel, mais une fonction phonostylistique importante (composante 6) et il caractérise le style des démonstrateurs, orateurs, et annonceurs (voir l'accent didactique de Lucci, 1979). Rossi (1993) assigne à l'ictus mélodique une fonction essentiellement rythmique (composante 3). La montée initiale (« initial rise ») a été interprétée comme la réalisation (facultative) d'une marque phonologique sous-jacente associée au début de chaque mot lexical en français (Vaissière, 1974, composante 4). Delattre (1961), puis Martin (1982) ont passé sous silence ce phénomène, le premier sans doute parce qu'il a analysé des interviews où cette montée est moins fréquente (Lucci, 1979), le second, parce ses corpus contiennent essentiellement des mots

courts, où la montée initiale est plus rare. La multiplicité des composantes a favorisé le foisonnement des interprétations et a empêché l'éclosion d'un modèle général. On citera néanmoins les efforts de Rossi (1993) pour un modèle théorique intégré ne comportant pas moins de six composantes. Ce modèle (incomplet, au regard des neuf composantes énoncés ci-dessus) n'a sans doute pas reçu toute l'attention qu'il méritait.

### 3. LES DIFFERENTES UNITES ET DOMAINES DE REALISATION DES MARQUES

Il n'y a naturellement pas « une » unité prosodique, mais un enchevêtrement d'unités.

#### 3.1. *Les unités prosodiques*

C'est d'abord, pour commencer par une des plus grandes unités, le paragraphe prosodique, marqué par une valeur de Fo extra haute vers son début, et une valeur extra basse à sa fin, avec une tendance générale pour les pics de Fo à décliner du début du paragraphe à sa fin (Lehiste, 1975). Le paragraphe est composé d'énoncés prosodiques, terminés par la succession d'une intonation montante sur l'avant dernier mot prosodique et par une chute finale de Fo (et d'intensité) sur le dernier mot. Un long énoncé peut être décomposé en groupes de souffle par une ou plusieurs pauses, respiratoires ou non, le groupe de souffle final se terminant par une chute finale de Fo (et d'intensité) et un allongement final moyen, les groupes internes par une montée de continuation et un fort allongement final. Un groupe de souffle est décomposé en syntagmes prosodiques commençant facultativement par un sommet de Fo et se terminant par une montée de continuation (le plus souvent) ou un simple pic de Fo, avec un allongement de la durée de la dernière rime. Les syntagmes sont à leur tour divisés en mots prosodiques, marqués par des fluctuations de Fo et un allongement final, et commençant éventuellement par une montée initiale (cf. la notion d'arc accentuel de Fonagy). La similarité des moyens acoustiques mis en œuvre à chacun des niveaux (renforcement initial, allongement final, déclinaison des valeurs de Fo) nous a conduit à proposer un modèle récursif superpositionnel (Vaissière, 1995), à contre courant des tendances actuelles (cf. l'extension à beaucoup de langues, le français compris, des pré-supposés sous-jacents au système de notation "Tobi", développé à l'origine pour l'anglais, Silverman et al, 1992).

#### 3.2. *Les domaines de réalisation des marques prosodiques : locaux et globaux*

Les marques prosodiques peuvent être très locales (infrasyllabiques) ou s'étendre sur un domaine plus vaste.

*Domaine infrasyllabique* : Donnons trois exemples. Dans le cas de la syllabe finale, l'allongement concerne essentiellement la rime (noyau plus coda), alors qu'il concerne la consonne d'attaque dans le cas de la syllabe initiale en français (Vaissière, 1988 pour une utilisation de cette propriété démarcative différenciée en RAP ; pour l'anglais, Campbell, 1992). Une position plus tardive de la cible de Fo dans la syllabe accentuée dans « Er hat **ge**logen » (allemand) contribue à signaler son caractère de nouveauté (Kohler, 1990). Les variations allophoniques fines (voir Straka, 1979, Fougeron et al, 1997 pour des études articulatoires) contribuent, dans une mesure qui reste à déterminer aux fonctions démarcatives et énonciatives.

*Domaine supra-lexical* : La continuité d'une tendance à décliner contribue à intégrer le paragraphe, l'énoncé et le syntagme. Son interruption à un temps donné marque une frontière. Sa diminution ou sa suppression sur un énoncé tout entier marque le caractère non terminal de cet énoncé (Thorsen, 1980). La diminution de la plage des variations de Fo et l'abaissement du registre sert à sa mise en retrait (parenthèse, incise). La durée est également régie globalement, du moins en partie. Les durées des syllabes moyennes dans un mot, un syntagme ou un énoncé, diminuent avec l'augmentation du nombre d'éléments qui les composent (Klatt, 1979 ; Bartkova et al, 1987). En d'autres termes, les mots ont tendance à être plus courts dans une phrase plus longue.

Phénomènes locaux et globaux coexistent et interagissent. Il a été prouvé que dans certains cas l'auditeur « normalise » les données locales par des données globales (par exemple, sur la perception de l'accent en fonction de la ligne de déclinaison, Pierrehumbert, 1979). De plus, l'importance relative des phénomènes locaux et globaux varie avec le degré de préplanification. Il y a moins de phénomènes globaux en parole spontanée, que dans la lecture soignée, la conférence, l'interview préparée (cf. Guaïtella, 1991). Ainsi, si des variations prosodiques subtiles en début de constituants permettent à l'auditeur de prédire la longueur de la phrase à venir (Grosjean, 1983), elles sont moins évidentes dans certains types de parole spontanée.

#### 4. APPLICATIONS TECHNOLOGIQUES

De nombreux secteurs technologiques ont identifié très tôt l'intérêt d'introduire la prosodie dans les systèmes automatiques. Par exemple, pour l'auditeur, il a été montré qu'une bonne prosodie aide à la mémorisation et l'accélération de la compréhension, au naturel de la phrase, au suivi de la conversation (effet cocktail), à la résistance aux distorsions et elle peut être strictement nécessaire pour lever certaines ambiguïtés syntaxiques. Son rôle apparaît donc comme majeur. Des efforts ont été faits depuis 1968 pour intégrer les connaissances prosodiques dans les systèmes automatiques. Faute de place, nous ne pourrions traiter que succinctement les domaines de la synthèse et de la reconnaissance.

##### 4.1. Synthèse et intégration des composantes syntaxique, rythmique, stylistique

La première "caricature" pour le français a été dérivée d'une voix d'un professionnel de la radio, à IBM (composantes intégrées : 1 et 3, Vaissière, 1971). D'autres imitations ont été créés à France-Télécom (voix publicitaire et sa propre voix, Emerard, 1977). Ces caricatures, qui transforment en règles explicites les tendances générales observées, avaient pour but de faire lire correctement une machine et il existe de très nombreux produits commerciaux. Des efforts récents, concrétisés par des thèses, visent à prendre en compte le contexte énonciatif, introduire des attitudes, changer de locuteur automatiquement, faire varier le débit, etc. L'approche est désormais essentiellement fondée sur l'utilisation de différentes méthodes statistiques sur de très larges bases de données et il faudra sans attendre le siècle prochain pour voir les premiers systèmes intégrés livrés au grand public.

#### 4.2. Reconnaissance

Les efforts pour intégrer les connaissances prosodiques en RAP remontent aux années 1970, avec les efforts de Wayne Lea pour l'anglais. Le premier but pour le français était d'utiliser la fonction démarcative de la prosodie dans les textes lus afin de compléter les hypothèses segmentales par des informations prosodiques à l'entrée du module lexical (dans le cadre de machines à dicter). Notre tâche s'est vite heurtée, entre autres, au problème non trivial de la séparation automatique entre "proéminences" de début (très fréquents en lecture) et de fin de mots (Vaissière, 1988). L'indéterminisme prosodique a empêché l'intégration de la prosodie dans les systèmes anciens et les systèmes actuels de RAP, par ailleurs très performants et fondés sur des méthodes statistiques. On assiste désormais à un renouvellement de l'intérêt pour l'utilisation de la prosodie pour la reconnaissance de la parole spontanée et du discours. Il existera sans doute bientôt des solutions partielles, adaptées à d'applications très ciblées (demande de billets d'avion, renseignements météo, etc.), où il est possible de construire un modèle statistique du comportement de l'utilisateur. Mais une solution générale n'est pas en vue.

#### 4. CONCLUSIONS

En France et à l'étranger, on assiste à un renouvellement de l'intérêt pour l'utilisation de la prosodie dans les systèmes automatiques, comme en témoigne le nombre croissant d'articles dans les séminaires et congrès réunissant essentiellement des ingénieurs, mais aussi des spécialistes de la communication parlée. Cet intérêt a été encouragé par le grand succès de l'application de méthodes statistiques en RAP (partie segmentale), et on espère que les statistiques permettront également de traiter l'aspect suprasegmental de la parole. On assiste aussi, comme en linguistique, à un recentrage de l'intérêt sur les phénomènes de parole spontanée et du dialogue. Les progrès technologiques dans le stockage et la vulgarisation des analyseurs ont favorisé la mise en place d'immenses bases de données de parole, prêtes à l'analyse statistique, mais toujours en attente d'étiquetage prosodique satisfaisant. Devant l'incapacité des linguistes (phonéticiens compris) de fournir des règles explicites (et pour cause



...) pour traiter des phénomènes prosodiques en parole spontanée, certains ingénieurs considèrent que la solution, si elle existe, passera par l'utilisation de statistiques sur d'immenses bases de données (systèmes de Markov, systèmes neuronaux, etc.) Le nombre de composantes prosodiques en jeu risque fort de compromettre les résultats en reconnaissance automatique de parole spontanée non contrainte. Les linguistes se doivent de participer à l'élaboration de paradigmes pour approcher le problème de l'intégration du plus grand nombre de composantes possibles en synthèse. Le temps est venu des modèles intégrant l'ensemble des connaissances déjà accumulées, qui seront le challenge du siècle prochain en prosodie.

## REFERENCES

- Atala, (1997), Prosodie et syntaxe, Volume 38, N° 1..
- Bartkova, K., C. Sorin, 1987, A model of segmental duration for speech synthesis in French. *Speech Communication*, 6, 245-260.
- Campbell, N., 1992, Syllabe-based segmental duration, in G. Bailly, Benoit, C. (ed), *Talking Machines: Theories, Models and Designs*. North-Holland. 211-224.
- Delattre, P., 1961, La leçon d'intonation de Simone de Beauvoir: Etude d'intonation déclarative comparée, *The French Review*, 35, : 59-67.
- Duez, D., Y Nishinuma, 1985, Le rythme en français: alternance des durées syllabiques, *Travaux de l'Institut de Phonétique d'Aix*, 10, 151-169.
- Emerard, F., 1977, Synthèse par diphtonges et traitement de la prosodie. Thèse de 3<sup>e</sup> cycle, Université de Grenoble 3.
- Foangy, I., (1983), *La vive voix*, Payot.
- Fougeron, C., S.-A. Jun, 1998, Rate Effects on French Intonation: Prosodic Organization and Phonetic Realization, *J. Phonetics*, 26 (1): 5-70.
- Fougeron, C., P. Keating, 1997, Articulatory lengthening at edges of prosodic domains, *J. Acoust. Soc. Am.*, 101 : 3728-3740.
- Grosjean, F., (1983), How long is the sentence : prediction and prosody in the on-line processing of language, *Linguistics* 21, 501-529.
- Guaïtella, I., 1991, *Rythme et parole: comparaison du rythme de la lecture oralisée et de la parole spontanée*, Aix-en-Provence, Thèse de Doctorat.
- Klatt, D. H., 1979, Synthesis by rule of segmental durations in English sentences, in B. Lindblom, Ohman, S. (eds), *Frontiers of Speech Communication Research*. New York and London, Academic Press. 287-300.
- Kohler, K., (1990), Macro and micro Fo in the synthesis of intonation, Paper in Lab. Phonology I, : *Between the grammar and physics of speech*, Kingston et al (eds.), Cambridge Un. Press, 115-138.
- Lehiste, I., 1975, The phonetic structure of paragraphs, in Cohen, A., Nootboom, S. (ed), *Structure and Process in Speech Perception*. Springer-Verlag.
- Léon, P., 1971, La joncture externe en français: nature et fonctions linguistiques et expressives, *Essai de Phonostylistique*, Studia Phonetica, Didier, 4 : 57-66.
- Lieberman, M., A. Prince, 1977, On stress and linguistic rhythm,, *Linguistic Inquiry*, 9 (2): 249-336.
- Lucci, V., 1979, L'accent didactique, Studia Phonetica, Ottawa: Didier, 15 : 107-121.
- Martin, P., 1982, Phonetic realisations of prosodic contours in French, *Speech Communication*, 1 : 283-294.

- Monnin, P., F. Grosjean, 1993, Les structures de performance en français: caractérisation et prédiction. *L'année Psychologique*, 93, 9-30.
- Pierrehumbert, J., (1979), The perception of fundamental frequency declinaison, *J. Acoust. Soc. Am.* 66, 363-369.
- Rossi, M., 1993, A model for predicting the prosody of spontaneous speech (PPSS model), *Speech Communication*, 13 : 87-107.
- Silverman, K., M. Beckman, et al., 1992, TOBI: A Standard for Labelling English Prosody. *Proceedings of the Second International Conference on Spoken Language Processing*, Bannf, Canada, 2, 867-870.
- Straka, G., 1979, *Les sons et les mots, choix d'études de phonétique et de linguistique*, Bibliothèque française et romane. Centre de Philologie et de littératures romanes de l'Université des Sciences Humaines de Strasbourg. Librairie C. Klincksieck.
- Thorsen, N., (1980), A study of the perception of sentence intonation – evidence from Danish, *J. Acoust. Soc. Am.*, 67, 1014-1030.
- Vaissière, J., 1971, *Contribution à la synthèse par règles du français*. Université des Langues et Lettres de Grenoble, Thèse de Troisième Cycle.
- Vaissière, J., 1974, On French prosody, *Quarterly Progress Report*, Massachusetts Institute of Technology, Research Laboratory of Electronics, 114 : 212- 223.
- Vaissière, J., 1975, Further Note on French Prosody, *Quarterly Progress Report*, Massachusetts Institute of Technology, Research Laboratory of Electronics, 115 : 251-261.
- Vaissière, J., 1983, Language-Independent Prosodic Features, in A. Cutler, R. Ladd (ed), *Prosody: Models and Measurements*. Berlin, Springer-Verlag. 53-66.
- Vaissière, J., 1988, The use of prosodic parameters in automatic speech recognition, dans Recent advances in speech understanding and dialog systems, Nieman et al (eds .), Springer-Verlag, 71-99.
- Vaissière, J., 1989, *Contribution à l'analyse des phénomènes de parole continue lue*, Thèse d'habilitation à diriger des recherches, Strasbourg.
- Vaissière, J., 1991, Rhythm, accentuation and final lengthening in French, in 108-120, *Music, Language, Speech and Brain*. Macmillan Press. 108-120.
- Vaissière, J., 1995, Phonetic explanations for prosodic cross-languages similarities. *Phonetica*, 52, 123-130.