

**CFA '18 LE HAVRE ■ 23-27 avril 2018**  
**14<sup>ème</sup> Congrès Français d'Acoustique**



**Le contrôle des instruments chanteurs**

C. d' Alessandro<sup>a</sup>, L. Feugère<sup>b</sup>, O. Perrotin<sup>c</sup>, S. Delalez<sup>d</sup> et B. Doval<sup>a</sup>

<sup>a</sup>Sorbonne Université, Centre National de la Recherche Scientifique, Institut Jean Le Rond d'Alembert,  
F-75005 Paris, France

<sup>b</sup>MIVEGEC, IRD, CNRS, Université de Montpellier, Montpellier, France

<sup>c</sup>School of Computing, University of Kent, Medway, UK

<sup>d</sup>Kepstral Audio, Toulouse, France

cda@dalembert.upmc.fr

La voix n'est pas un instrument de musique, car il n'y a pas d'objet externe mis en jeu par les membres ou par le souffle. La synthèse numérique permet pour la première fois une coupure entre le sujet et sa voix, en construisant des instruments chanteurs manipulés par les mains, les pieds, ou par toutes sortes d'interfaces humain-machine. Cependant, les possibilités de contrôle des instruments chanteurs sont encore limitées à certains aspects, car une transposition des gestes interne du chant effectués par l'appareil vocal en gestes externes des membres ne va pas de soi. Certains gestes sont analogues, d'autres sont médiatisés par un espace perceptif. Les travaux menés sur trois instruments chanteurs sont présentés: le contrôle de l'intonation par un stylet sur une tablette graphique et des gestes d'écriture ; le contrôle des voyelles et de la qualité vocale sur une surface ; le contrôle bimanuel de l'articulation consonantique; le contrôle syllabique du rythme. Les modèles de synthèse sous-jacents utilisent soit la simulation du modèle source filtre, soit la modification d'échantillons préenregistrés et étiquetés. Le contrôle des instruments chanteurs est multimodal, impliquant l'ouïe ainsi que la vue, le toucher, et la kinesthésie. Cette combinaison sensorielle et motrice permet dans certains cas de rendre les instruments chanteurs plus justes et précis que la voix, la vue privilégiant les aspects mélodiques et l'audition des aspects rythmiques. Miroir de la voix, l'instrument chanteur autorise toutes sortes de spéculations : musicales bien sûr, avec le Chorus Digitalis, chœur de voix de synthèse, mais aussi pour l'analyse des pratiques vocales, pour l'éducation ou la rééducation, en renforçant l'apprentissage de gestes vocaux par des traces visuelles, des gestes manuels ou corporels. Le statut symbolique de la voix est également affecté par la possibilité de contrôler et de produire le son vocal en dehors du corps : corps augmenté, mise en scène de l'expression vocale, double de la voix, jeu de la voix d'un.e autre.

# 1 Introduction

## 1.1 Les avantages de l'externalisation

La voix n'est pas un "instrument" de musique, au sens d'un artefact mis en vibration par les membres ou par le souffle. Les organes vocaux sont internes, en grande partie invisibles, et contrôlés de façon complexe par plusieurs ensembles musculaires (respiration, phonation, articulation). Le contrôle vocal est donc, par nature intéroceptif, alors qu'il est davantage kinesthésique et extéroceptif pour les instruments de musique. Ce contrôle est bien sûr auditif dans les deux cas, mais la modalité visuelle ne joue aucun rôle pour la voix.

L'avènement de la synthèse numérique permet pour la première le rendu d'un son indéniablement vocal par un dispositif instrumental externe, mis à distance de l'appareil vocal. Les "instruments chanteurs" sont "manoeuvrés" par les mains, les pieds, ou par toutes sortes de capteurs ou d'interfaces humain-machine. Cette mise à distance pose la question du contrôle du chant dans des termes tout à fait différents de ceux du contrôle d'un instrument acoustique. Il faut en effet s'abstraire de l'espace physique, utiliser des espaces perceptifs intermédiaires et mettre en place des espaces de contrôle [5]. Les avancées sont suffisantes pour présenter aujourd'hui des instruments chanteurs effectivement joués en concert, qui ouvrent également des perspectives pédagogiques, voire thérapeutiques.

## 1.2 Milieu sémiotique du chant: une triple tâche

Le chant est le point de rencontre entre organe vocal, langage, musique et expression. Langage et musique sont des systèmes de signes (grammaire et soifège) culturellement déterminés, qui peuvent aussi exister en l'absence de la voix (musique instrumentale, langage gestuel). Ces signes sont des moyens de représentation de ce qui est dit ou joué, ils portent la signification du texte et la partition.

Mais la voix seule, évènement vocal pur, exprimée sans parole ou sans référence à un système musical est déjà du chant. La musicalité, ou l'expression, est proche du geste et du mouvement corporel, en suivant la célèbre formule

"la musique est la science des mouvements bien ordonnés" (*musica est scientia bene movendi*) [6], ce qui reflète bien l'étymologie du mot émotion.

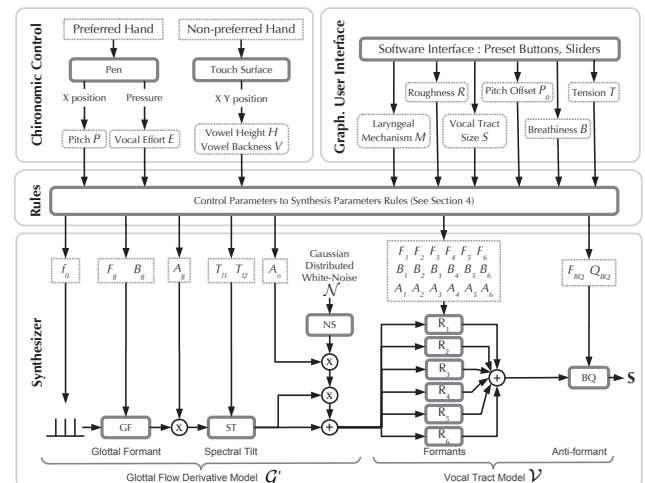
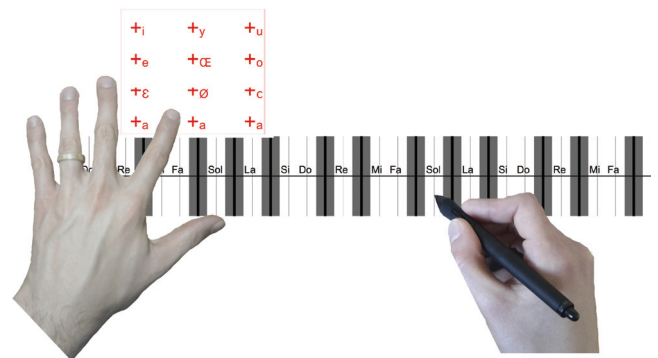


Figure 1: Cantor Digitalis. En haut: principe du contrôle bimanuel des voyelles et de l'intonation. En bas: schéma du synthétiseur paramétrique. Figures extraites de [12]

La variation expressive qui peut-être symbolique (attitudes, affects typiques) ou non, résiste à la description structurale. L'expression, toujours située et incarnée relève de l'interprétation (herméneutique), des moyens de séduire ou convaincre (rhétorique), en bref de la manière de dire ou de jouer. L'expression est essentielle en situation de

communication pour donner le sens musical, au delà de la signification du texte. Les gestes du chant effectuent donc plusieurs tâches simultanément, suivant la triple contrainte linguistique, solfégique et expressive:

**tâche linguistique** : l'unité minimale pour rendre le chant intelligible, semble être la syllabe. En effet, le phonème est une abstraction linguistique, mais la réalisation effective d'un geste vocal est toujours une syllabe. Sur et par la syllabe peuvent se construire la double articulation phonème/mot, la prosodie, et les niveaux linguistiques supérieurs. La contrainte linguistique est donc de produire des syllabes intelligibles, éventuellement par combinaison de phonèmes.

**tâche solfégique** : un solfège, explicite et formalisé ou non, existe pour la plupart des langages musicaux traditionnels. L'unité minimale est alors la note, soit une hauteur mélodique (suivant une échelle mélodique), une valeur de durée, éventuellement des nuances d'intensité ou le timbre. La contrainte solfégique est donc de produire des hauteurs précises avec des durées contrôlées.

**tâche expressive** : l'expressivité relève principalement de la variation, du phonostyle, par rapport aux contraintes linguistique et musicale. Cette variation se manifeste à toutes les échelles (micro-mélodie, ornements, autant que mouvement sur une phrase entière), en particulier dans la qualité vocale.

Il faut nuancer ce tableau, car le chant n'a pas forcément besoin d'être intelligible: scat, glossolalies, langues inconnues ou imaginaires, proto-langue comme dans le cas du "human beatbox" qui est constitué d'un lexique de symboles sonores. Un solfège n'est nullement obligatoire, en distinguant les "musique de notes" qui s'appuient sur un solfège, des "musiques de son" qui peuvent volontiers s'en passer.

Afin de pouvoir évaluer les résultats obtenus, il nous a semblé important dans le développement des instruments chanteurs de maintenir un paradigme d'imitation, en étant capable de jouer des notes de musiques et des paroles intelligibles. L'imitation est la meilleure épreuve de la qualité sonore et de la finesse de contrôle. Mais bien sûr, pour la création musicale, le modèle peut être étendu en offrant de nouvelles possibilités de contrôle et des gestes et des sons qui se démarquent de l'imitation vocale.

### 1.3 Le geste vocal

Dans l'espace physique, le geste vocal comprend des gestes respiratoires (thorax, abdomen), des gestes laryngés, des gestes articulatoires (lèvres, voile du palais, langue, mâchoires). Certains gestes co-verbaux sont sans effet sonore, gestes oro-faciaux non articulatoires (yeux, sourcils), gestes des membres et du corps. Les gestes produisant du son se divisent en :

**geste articulatoire**: le geste articulatoire correspond à une fermeture/ouverture plus ou moins complète dans le conduit vocal, depuis la glotte jusqu'au lèvres. L'intéroception de ce geste est une contraction, ou un mouvement des articulatoires, coordonné avec

le contrôle du souffle par contraction abdominale ou thoracique, et l'activité glottique. Le geste articulatoire est rythmé par l'oscillation mandibulaire.

**geste phonatoire**: le geste intonatif, mélodique ou phonatoire, correspond à l'abduction/adduction et à la tension du larynx, mis en jeu pas un ensemble complexe de muscles et cartilages coordonnés avec la respiration. L'intéroception de ce geste est une contraction, ainsi que le contrôle du souffle par contraction abdominale ou thoracique.

Le rythme vocal est construit sur la syllabe, qui correspond au geste articulatoire d'ouverture/fermeture de la mandibule, et sur le geste phonatoire et respiratoire. Le geste expressif est également porté par le geste articulatoire et/ou le geste phonatoire : il correspond à la conduite et à la variation des des mouvements articulatoires et laryngés. La plupart des instruments chanteurs ne prennent pas en compte la respiration, à l'exception de Skeezebox [8] qui utilise un soufflet d'accordéon.

## 2 Cantor Digitalis

Cantor Digitalis est un instrument chanteur construit sur un synthétiseur vocal paramétrique [12]. Les gestes de contrôle sur une tablette graphique utilisent les deux mains: main préférée avec un stylet, et la main non-préférée avec les doigts. La synthèse paramétrique permet de calculer les formants, l'intonation, et un modèle de la source vocale. L'instrument produit des voyelles (voisées ou non) et toutes sortes de sons dérivées du modèle vocalique, associés à de multiples gestes expressifs.

### 2.1 Ecriture mélodique

La chironomie, ou externalisation par les mouvements de la main du geste phonatoire passe par la construction d'un espace perceptif, l'espace mélodique. Cette espace perceptif, à une dimension, peut se déployer dans un espace à 3 dimensions (geste libre), à deux dimensions ou à une dimension. Le geste chironomique (spatial) diffère beaucoup du geste laryngé (tension glottique).

Le geste d'écriture semble un excellent compromis entre liberté et facilité d'apprentissage [2, 17]: le clavier (même PME, Multi Polyphonic Expression) favorisant les notes discrètes, et la précision étant difficile avec le geste libre (type Theremin) [1].

Dans le Cantor Digitalis c'est l'image d'un clavier qui représente la hauteur mélodique (voir l'architecture de l'instrument sur la Figure 1). Le geste d'écriture semble particulièrement facile à utiliser pour le contrôle mélodique. En effet, c'est un geste acquis dès l'enfance, à l'aide de la main préférée.

La justesse et la précision du contrôle mélodique par un stylet sont meilleures qu'avec la voix [4]. L'audition, la vision, la kinesthésie coopèrent dans cette tâche gestuelle de visée de cible, alors que le chant naturel ne peut compter que sur l'intéroception et l'audition [16].

### 2.2 Le geste vocalique

Le contrôle physique des voyelles passe par la conformation du conduit vocal qui dépend de la position

des articulateurs. Pour le contrôle temps-réel, les voyelles peuvent se répartir dans un espace perceptif décrit par 2 dimensions acoustiques, le triangle vocalique, défini par la fréquence centrale des deux premiers formants spectraux. C'est ce mode de contrôle qui est proposé dans le Cantor Digitalis à l'aide de la main non-préférée. Les transitions entre voyelles sont assez réalistes, même si l'espace est acoustique plutôt qu'articulatoire. Un ensemble de voyelles cibles permet de reconstituer par interpolation tout l'espace vocalique. Si on le souhaite, les deux dimensions de l'espace vocalique peuvent se projeter sur un seul axe de la tablette, en fixant un parcours particulier des voyelles dans le plan.

### 2.3 La qualité vocale

La variation de qualité vocale peut être réduite à deux dimensions principales, l'effort et la tension. L'effort vocal se manifeste par la pression subglottique, la résistance glottique, la contraction supraglottique. La tension et le relâchement de la voix résulte de gestes d'adduction/abduction des plis vocaux. En passant par un espace perceptif, le contrôle gestuel de l'effort ou de la tension peut être associé à une dimension spatiale (par exemple verticale sur la tablette) ou bien, de façon plus analogue au geste physique, à la pression sur le stylet. Un autre aspect de la qualité vocale est l'allongement du conduit vocal, dont la perception peut être contrôlé dans un espace à une dimension, par exemple sur un des axes de la tablette graphique.

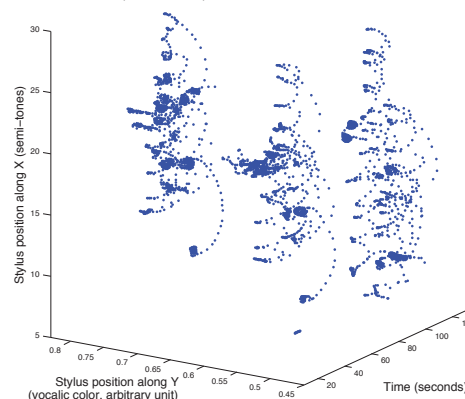
### 2.4 L'expression: un exemple choral

Les figures 2 et 3 montrent les tracés graphiques et la partition d'un choral à 4 voix, chanté par le Chorus Digitalis. Le choral est chanté sur les voyelles /a/ et /u/, représentées sur l'axe Y, avec le temps sur l'axe X et la hauteur mélodique sur l'axe Z. La première phrase du choral est reprise en /a/, puis en /u/ ce qui correspond aux deux premiers ensembles, puis la seconde phrase alterne des /a/ et /u/ et continue sur /u/, ce qui donne le 3e ensemble de notes. Le style gestuel (et donc l'expression) de chaque musicien ne ressort nettement des tracés, alors que les contenus mélodique et rythmique sont comparables pour les 4 voix. On observera: le vibrato, la forme et le netteté des transitions entre notes, la justesse, la simplicité ou la complexité des gestes. Cependant, il faut noter que l'effort vocal n'est pas rendu sur la figure, ce qui serait nécessaire pour préciser l'expression.

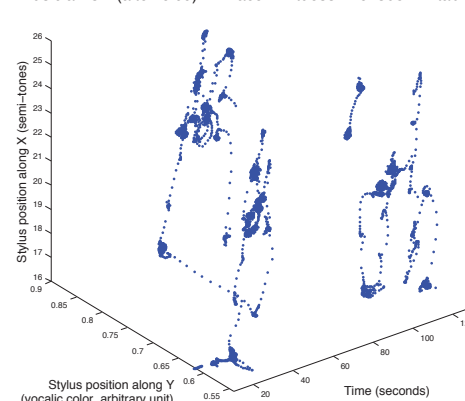
## 3 Digitartic

La durée du geste consonantique est en général plus brève que celle du geste vocalique dans la syllabe. La dimension de l'espace consonantique est difficile à évaluer. On distingue au moins les dimensions de lieu d'articulation (labial, dental, palatal, vélaire etc.), de mode d'articulation (plosif, fricatif, nasal, liquide), de voisement (voisé, non voisé). En plus de ces dimensions intervient la coarticulation avec les voyelles, soit deux dimensions supplémentaires, et les différentes phases temporelles pour certaines consonnes (phases très courte, quelque dizaines voire quelques millisecondes). Le contrôle direct du geste consonantique pour tous les phonèmes semble problématique. Une extension du Cantor Digitalis, l'instrument Digitartic (voir l'architecture sur la

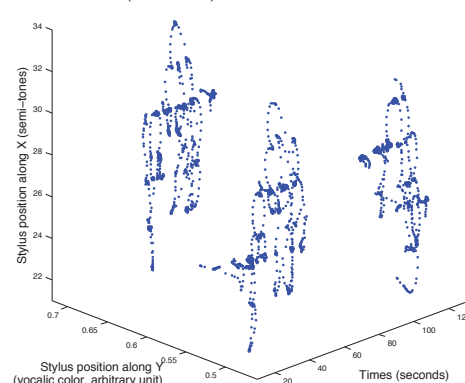
musician BD (bass voice) – Wacom Intuos4 224x140mm tablet



musician CA (alto voice) – Wacom Intuos3 229x305mm tablet



musician HM (tenor voice) – Wacom Intuos3 229x305mm tablet



musician LF (soprano voice) – Wacom Intuos4 224x140mm table

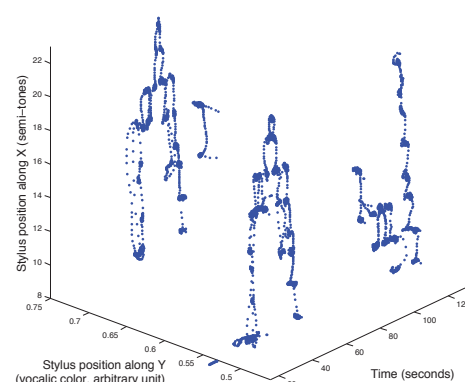


Figure 2: Mouvements mélodiques et vocaliques pour les 4 parties du choral 3.





Figure 3: Choral "Wie schön leuchtet der Morgenstern", harmonisation J.S. Bach

Figure 4), synthétiseur paramétrique par règles [13] permet le contrôle gestuel direct d'un sous ensemble de consonnes. Il est possible d'articuler des syllabes CV, VC, CVC, mais pas de restituer un énoncé tant soit peu complexe. A l'usage, l'instrument s'est révélé difficile à maîtriser, le rendu de consonnes intelligibles étant difficile.

L'instrument Digitartic montre les limites de la synthèse paramétrique pour la parole articulée, lorsque l'on cherche à contrôler de façon synchrone les variations rapides et détaillées de nombreux paramètres. D'une part l'espace des paramètres consonantiques a beaucoup de dimensions (GloveTalk, utilisait les 5 doigts, le geste manuel en 3D et les pieds afin de multiplier les dimensions contrôlées [11]), d'autre part la dynamique des mouvements consonantique est beaucoup plus rapide que celle du geste d'écriture.

## 4 Vokinesis

### 4.1 L'organisation du geste syllabique

Pour s'affranchir des problèmes de contrôle de la synthèse paramétrique, la synthèse par échantillonnage consiste non plus à calculer les paramètres d'un modèle de production du signal, mais à modifier des échantillons pré-enregistrés de signal.

Des systèmes de modification temps-réel de la fréquence fondamentale et du rythme ont été développés. Les résultats montrent un contrôle mélodique excellent puisque pour l'imitation mélodique, le geste est perceptivement équivalent à la voix [3].

Le rythme de séquences vocales pré-enregistrées peut être contrôlé simplement en variant la vitesse de lecture, ou le débit. Il ne s'agit pas d'un contrôle à proprement parler rythmique, mais plutôt d'un contrôle du défilement du temps. Une autre façon de contrôler le temps est de piloter le défilement des échantillons (le pointeur temporel), ce qui permet aussi de lire le signal à l'envers ou de jouer à volonté n'importe quelle partie de la séquence. Ici le rythme produit correspond au mouvement rythmique du pointeur, mais pas forcément au rythme syllabique.

### 4.2 Contrôle du rythme syllabique

Contrôler le geste syllabique fait appel à la phonologie articulaire. La syllabe comprend trois éléments: l'attaque (optionnelle, une ou plusieurs consonnes), le noyau vocalique (obligatoire), et la coda (optionnelle, une ou plusieurs consonnes). En suivant la théorie frame-content [14], le mouvement syllabique est constitué d'un cadre vocalique (frame) qui va de voyelle à voyelle, et qui correspond aux mouvements articulatoires de grande amplitude (ouverture de la mandibule, avancée de la langue) et d'un contenu, mouvements consonantiques plus rapides qui s'inscrivent dans le cadre.

Le geste syllabique implique donc au moins deux grands sous-ensembles, les gestes vocaliques et les gestes consonantiques, qui s'inscrivent dans le mouvement rythmique des syllabes. Notons que cette théorie est proche de la description classique du rythme comme successions d'arsis (levée du pied) et thesis (pose du pied). Pour le contrôle du rythme syllabique, il faut au préalable repérer la position des phonèmes dans le signal. Jouer un rythme revient à viser des cibles dans les syllabes. Ce mouvement

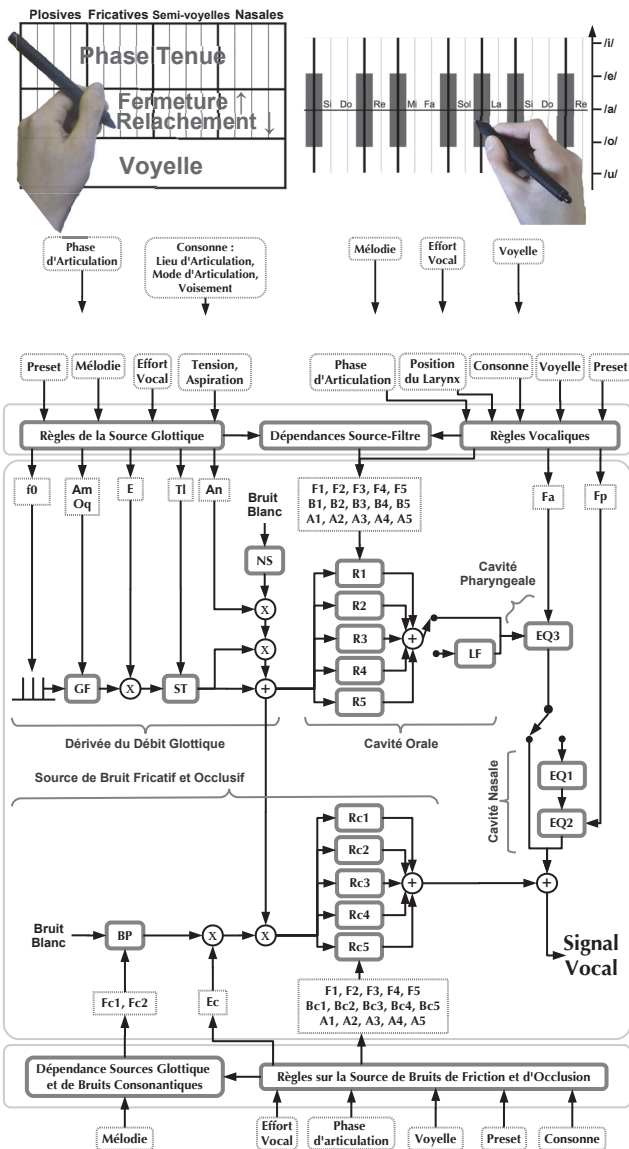


Figure 4: Digitartic. En haut: principe du contrôle bimanuel des consonnes et de l'intonation. En bas: schéma du synthétiseur paramétrique. Figures extraites de [13]

comprend deux phases pour chaque syllabe de la même façon qu'une note de musique comprend deux phases: la phase d'attaque (onset) et la phase d'extinction (offset). La première phase est "l'ouverture" de la syllabe, entre le point consonantique le plus fermé et le point de la voyelle correspondant à la pleine ouverture de la mandibule: cette phase inclut l'attaque et le début du noyau vocalique, c'est la posée de la syllabe ou thesis. La seconde phase, de fermeture, va de l'ouverture de la mandibule dans le noyau vocalique à la fermeture dans la coda, c'est la transition, la levée de la syllabe ou arsis. Les deux phases peuvent être contrôlées par divers dispositifs, un bouton on/off ou des contrôleurs continus, comme potentiomètres ou pédales.

Les geste utilisés sont semblables aux gestes de battue rythmique avec les mains ou avec les pieds. Il a une analogie entre les mouvements rythmique des membres et de la mandibule. Les gestes d'écriture sur une tablette graphique sont utilisés pour le contrôle mélodique. L'architecture du système Vokinesis est décrite dans la Figure 5.

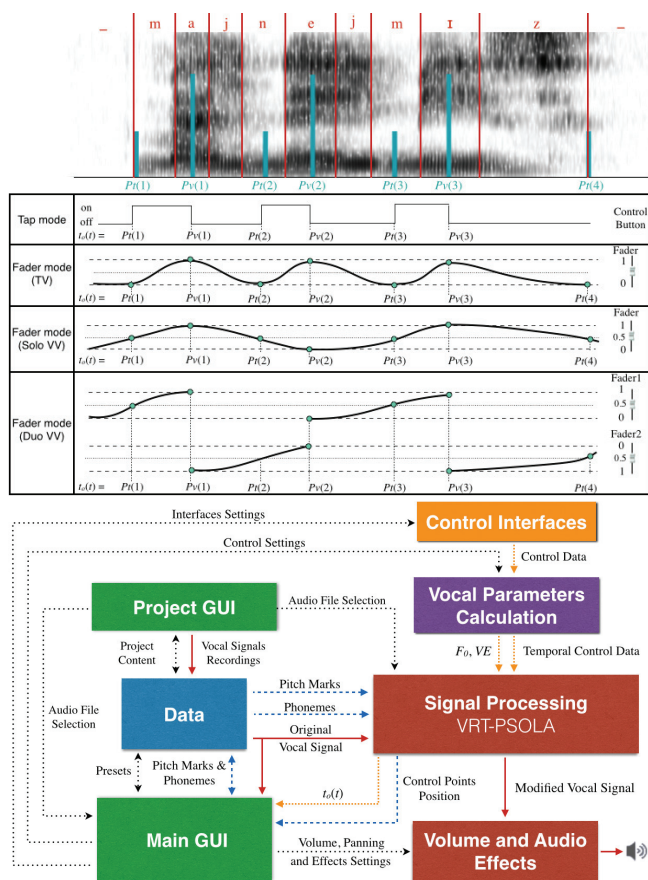


Figure 5: Vokinesis: en haut, étiquetage du signal en arsis et thesis, et contrôle rythmique par un bouton, un ou plusieurs potentiomètres; en bas, architecture du système. Figures extraites de [9, 10]

## 5 Qui chante?

Le contrôle de la synthèse numérique de la voix se démarque de celui des instruments de musique. Trois niveaux de transposition gestuelle se dessinent: au premier niveau, celui du piano numérique ou du saxophone MIDI, l'interface de contrôle de l'instrument acoustique est conservée; au second niveau, celui de la modélisation

physique des instruments monophoniques, comme les flûtes, les gestes de contrôle peuvent être transposés sur des contrôleurs MPE [7] (EigenHarp, Haken Continuum, Seaboard Roli, LinnStrument etc.); au troisième niveau, celui de la voix, des geste linguistiques s'ajoutent au gestes solfégiques et expressifs. L'externalisation des gestes de contrôle vocal fait appel à plusieurs niveaux d'abstraction, résumés dans le Tableau 1. Le geste de contrôle peut être analogue au geste vocal: dans le cas du contrôle rythmique, la transposition manuelle ou pédestre du mouvement oscillatoire de la mandibule fonctionne bien, est facile à apprendre, et permet une grande précision; la résistance du larynx et la pression subglottique, responsables de l'effort vocal, sont bien rendus par a pression du stylet sur la tablette.

Pour les geste mélodiques, un niveau d'abstraction supplémentaire est nécessaire. Il faut passer par l'espace perceptif des hauteurs musicales, et projeter cet espace sur, par exemple, une surface. D'autre gestes, comme l'allongement du conduit vocal passent par un espace perceptif pour être rendus par des gestes qui n'ont que peu d'analogie avec le geste physique.

Enfin les gestes articulatoires, surtout les consonnes semblent beaucoup trop rapides pour être contrôlés dans le détail. Il est possible de contrôler la lecture des échantillons, mais difficile actuellement de réaliser à volonté tout geste consonantique dans la temporalité de la synthèse.

Les gestes employés utilisent le retour haptique, du stylet sur une surface, des boutons ou des pédales. Comme pour les instruments acoustiques, la résistance opposée par des dispositifs physiques permet de caler les gestes et d'affiner leur précision. Le contrôle des instruments chanteurs est multimodal, impliquant l'ouïe, la vue, le toucher et la kinesthésie. Cette combinaison sensorielle et motrice permet dans certain cas de rendre les instruments chanteurs plus justes et plus précis que la voix. Il est possible d'ajouter une automatisation ou semi-automatisation du contrôle vocal: correction de justesse pour la mélodie [15], durée des transitions voyelles/consonnes par exemple.

Outre une réflexion sur la nature du contrôle gestuel, Les instruments chanteurs offrent des possibilités musicales nouvelles. L'instrumentiste devient le chanteur d'une "voix", mais c'est une voix artificielle, qui n'est pas la sienne, aux possibilités extrêmes en ce qui concerne le registre, le timbre, la vélocité voire même le sens du temps. Les instruments autorisent l'imitation, jouer une voix qui ressemble à une voix (au système de reproduction sur haut-parleurs près), et l'extension de la voix. Cette extension provient de paramètres ou de gestes impossibles dans le cas naturel.

Externaliser la voix la désincarne, ou l'incarne différemment. Le statut symbolique de la voix est questionné, voire affecté par la possibilité de contrôler et de produire le son vocal en dehors du corps : corps augmenté, mise en scène de l'expression vocale, double de la voix, jeu de la voix d'un.e autre. Cette coupure des représentations (sociales, corporelles, musicales) ouvre vers je jeu, l'apprentissage, l'analyse, une écoute différente. L'étrangeté de ce double artificiel de la voix peut jouer dans plusieurs sens: il dérange et attire, amuse et rebute.

Les instruments chanteurs sont susceptibles d'applications pédagogiques. L'instrument interne devenu externe, la voix qui s'expose permet l'analyse des pratiques vocales, pour l'éducation ou la rééducation. Tout comme le renforcement

TABLEAU 1: Gestes dans les espaces corporel, perçus et de contrôle instrumental. Instruments: C Cantor Digitalis, D Digitartic, V Vokinesis. Tâches: L Linguistique, S Solfégique, E expressive.

	espace corporel	espace perceptif	espace instrumental	instrument	tâche
phonation	tension plis vocaux résistance glottique, pression résistance glottique	mélodie effort vocal tension vocale	stylet, axe ou surface pression stylet, axe pression stylet, axe	C D V C D V C D	L S E E S E
articulation	longueur conduit vocal forme conduit vocal lieu d'articulation mode d'articulation	taille perçue voyelle consonne consonne	stylet axe doigt surface / stylet axe stylet, surface dynamique stylet, espace	C D C D D D	E L L L
rythme	oscillation mandibulaire	arsis/thesis	bouton/pédales	V	S E

multimodal permet une excellente justesse et précision, il ouvre des perspectives pour l'apprentissage des gestes vocaux en combinant perception et production de traces visuelles, de gestes manuels ou corporels et de sons vocaux.

## References

- [1] C. d'Alessandro, N. D'Alessandro, S. Le Beux, J. Simko, F. Çetin, Hannes Pirker, "The Speech Conductor: Gestural Control of Speech Synthesis" proceedings of eINTERFACE 2005 The Summer Workshop on Multimodal Interfaces, Mons, Belgium, pp. 52-61.
- [2] N. D'Alessandro, C. d'Alessandro, S. Le Beux, B. Doval "Real-time CALM Synthesizer New Approaches in Hands-Controlled Voice Synthesis", Proceedings NIME06, Paris, France, pp. 266-271.
- [3] Christophe d'Alessandro, Albert Rilliard, and Sylvain Le Beux "Chironomic stylization of intonation" J. Acoust. Soc. Am., 129(3), march 2011, 1594-1604.
- [4] Christophe d'Alessandro, Lionel Feugère, Sylvain Le Beux, Olivier Perrotin, and Albert Rilliard (2014) , Drawing melodies : evaluation of chironomic singing synthesis, J. Acoust. Soc. Am. 135 (6), 3601-3612.
- [5] Arfib, D., Couturier, J., Kessous, L., Verfaillie, V. (2002). *Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. Organised Sound*, 7(2), 127-144. doi:10.1017/S1355771802002054
- [6] Augustin d'Hippone *De Musica (Traité de la musique)*, Livre 1, Chap. III, ca 386-387.
- [7] Geert Bevin, MIDI specifications for Multidimensional Polyphonic Expression (MPE), <http://expressiveness.org/>, consulté le 21 mars 2018.
- [8] Cook, P., Leider, C., *SqueezeVox: A New Controller for Vocal Synthesis Models*, Proceedings ICMC00, Berlin, 2000.
- [9] Samuel Delalez, Christophe d'Alessandro " Vokinesis : syllabic control points for performative singing synthesis", Proceedings of NIME17,2017, Copenhagen, Denmark 198-203.
- [10] Samuel Delalez, Christophe d'Alessandro " Adjusting the Frame: Biphasic Performative Control of Speech Rhythm ", Proc. INTERSPEECH 2017, Stockholm, Sweden, August, 864-868.
- [11] Sidney Fels et Geoffrey Hinton : Glove-talk ii - a neural-network interface which maps gestures to parallel formant speech synthesizer controls. IEEE Transactions on Neural Networks, 9(1):205-212, January 1998.
- [12] Lionel Feugère, Christophe d'Alessandro, Boris Doval, Olivier Perrotin, " Cantor Digitalis: Chironomic Parametric Synthesis of Singing ", Eurasip J. Audio Speech Music Processing, 2017.
- [13] Lionel Feugère, Christophe d'Alessandro " Synthèse à partir du geste de la voix chantée: instruments Cantor Digitalis et Digitartic", *Traitement du Signal*, 32 (4), décembre 2015, 417-442
- [14] P. F. MacNeilage, The frame/content theory of evolution of speech production Behavioral and brain sciences, vol. 21, no. 04, pp. 499-511, 1998.
- [15] Olivier Perrorin, Christophe d'Alessandro, "Dynamic Pitch Warping for Intonation Correction of Digital Musical Instruments " , ACM Transactions on Computer-Human Interaction (TOCHI) 23(3), June 2016, Article No. 17,
- [16] Olivier Perrotin, Christophe d'Alessandro, "Seeing, listening, drawing: interferences between sensorimotor modalities in the use of a tablet musical interface", ACM Transactions on Applied Perception, 14(2), October 2016.
- [17] Marcelo M.Wanderley, Jean-Philippe Viollet, F. Isart et Xavier Rodet : On the choice of transducer technologies for specific musical functions. In Proceedings of the International Computer Music Conference (ICMC), volume 2000, Berlin, Germany, September 2000.