

# QUEL AJUSTEMENT DE HAUTEUR MÉLODIQUE POUR LES INSTRUMENTS DE MUSIQUE NUMÉRIQUES ?

Olivier Perrotin  
LIMSI-CNRS  
olivier.perrotin@limsi.fr

Christophe d'Alessandro  
LIMSI-CNRS  
cda@limsi.fr

## RÉSUMÉ

Deux méthodes d'ajustement de justesse mélodique sont présentées et comparées en terme de justesse et d'expressivité : l'ajustement *notes étendues* et l'ajustement *élastique*. Le but de ces ajustements est d'améliorer la justesse du musicien, tout en minimisant les contraintes nuisant à son expressivité. Les méthodes diffèrent par le type de transformation appliquée à la hauteur jouée. De plus, leurs dynamiques d'application sont paramétrables. La comparaison des deux montre des performances similaires en termes de justesse. En revanche la limitation de l'expressivité dépend grandement de la dynamique d'application choisie. Finalement, l'ajustement *élastique* permet la plus grande expressivité.

## 1. INTRODUCTION

### 1.1. Justesse vs. expressivité

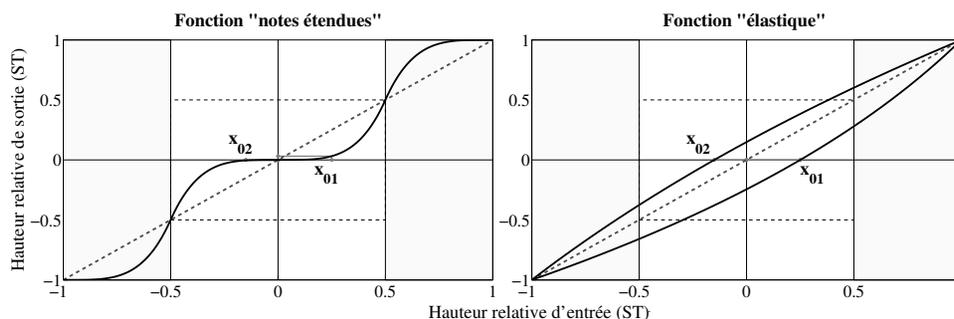
La plupart des styles de musiques sont construits sur des gammes proposant un ensemble fini de notes. Il est indispensable pour le musicien de pouvoir atteindre ces notes précisément car la moindre erreur détériore grandement le rendu de la prestation, quelle que soit la qualité sonore de l'instrument. Par ailleurs, il est d'usage d'agrémenter la trajectoire mélodique de modulations reflétant une des formes d'expression possibles du musicien [8], qu'on appellera par la suite expressivité mélodique ou expressivité. Trois modulations de trajectoires communément utilisées sont le *portamento*, transition continue et courte entre deux notes, le *glissando*, transition continue et longue entre deux notes, et le *vibrato*, oscillation autour d'une hauteur stable. Savoir combiner justesse et expressivité mélodiques est une étape importante dans l'apprentissage d'un instrument.

Cette dualité justesse vs. expressivité est corrélée à l'espace de contrôle de la hauteur mélodique. Les instruments de musique acoustiques mélodiques sont catégorisés en trois espaces de contrôle. Certains possèdent un espace entièrement discret (p. ex. claviers). Seules les notes d'une échelle donnée sont proposées, induisant une justesse parfaite mais sans modulation possible. A l'inverse, d'autres instruments présentent un espace de contrôle mélodique continu (p. ex.

cordes frottées, voix). Atteindre une note juste est plus difficile car la moindre déviation est perçue comme une fausse note. En revanche, toute modulation de hauteur est possible, favorisant l'expressivité mélodique du musicien. La troisième catégorie possède à la fois une échelle discrète proposée par un certain doigté, et un contrôle continu local réalisé à l'embouchure (p. ex. instruments à clés ou à pistons). L'identité d'un instrument est en partie définie par son espace de contrôle mélodique. Par conséquent le choix de ce dernier est inévitable lors de la construction d'un nouvel instrument de musique numérique mélodique. Il est d'abord possible de concevoir un espace de contrôle discret, tout en permettant des variations de hauteur expressives autour des notes cibles, p. ex. molette "Pitch bend" sur claviers (Minimoog, Yamaha DX7), à l'image de la troisième catégorie d'instruments. Le problème inverse est traité dans ce papier : la conception d'un espace de contrôle continu augmenté d'un ajustement de hauteur mélodique facilitant la justesse.

Ces recherches découlent de l'utilisation de la tablette graphique comme contrôleur d'instruments de musique numériques [18]. Les hautes résolutions spatiale et temporelle des tablettes d'aujourd'hui, ainsi que la familiarité des gestes associés (écriture, dessin) font de cette interface un candidat de choix pour un contrôle expressif de la hauteur mélodique de l'instrument. Différentes équipes de recherches ont montré un intérêt dans cette interface pour le contrôle de la voix chantée [16],[9],[1],[2],[11]. De plus, l'évaluation d'un tel type de contrôle a montré qu'il permet de jouer très juste en situation expérimentale [3],[4]. Dans des conditions de jeu réelles, la tâche est néanmoins plus difficile. C'est pourquoi nous proposons deux méthodes d'ajustement de justesse afin de faciliter le jeu d'instruments numériques à espace de contrôle continu. Le but de ce travail est de comparer les différentes méthodes d'ajustement en termes d'amélioration de justesse et de préservation de l'expressivité mélodique de l'instrumentiste.

Après une revue des méthodes d'ajustement existantes en section 1.2, ces deux méthodes sont présentées en partie 2 et comparées en section 3. Les résultats obtenus sont discutés en dernière partie.



**Figure 1.** Fonctions de relation entre hauteurs d'entrée et de sortie exprimées en demi-tons (ST) relativement à la note cible (point (0,0)). Gauche : ajustement *notes étendues* ; Droite : ajustements *élastiques*, une fonction est calculée pour chaque position d'entrée.

## 1.2. Méthodes d'ajustement existantes

Beaucoup de constructeurs d'instruments de musique numériques à espace de contrôle continu ont intégré des ajustements de justesse pour améliorer le confort du musicien. C'est le cas d'instruments tels que le Haken Continuum [6], le Seaboard [10], le LinnStrument<sup>1</sup>, les applications iOS Morphwiz<sup>2</sup> et Garageband<sup>3</sup>, le piano augmenté de McPherson [12] ou le *Cantor Digitalis* [14].

L'ajustement de justesse consiste à découpler la *hauteur d'entrée* jouée par le musicien sur l'interface de la *hauteur de sortie* liée au synthétiseur et perçue par le musicien. Bien que l'implémentation des méthodes d'ajustement de certains produits commerciaux ne soient pas explicitées, deux catégories d'ajustement sont identifiées. La première utilisée par Haken [7] est une translation de la hauteur d'entrée vers la note cible. A chaque intervalle de temps  $\Delta TS$ , la trajectoire d'entrée est corrigée d'un pas de correction fixe  $\pm \Delta CS$  en demi-tons. Cette méthode ajuste la valeur moyenne locale de la trajectoire sans toucher aux variations relatives à l'expressivité.

La deuxième catégorie d'ajustements introduit une fonction de correspondance entre hauteurs d'entrée et de sortie. Celle-ci peut être d'abord définie de manière statique, i.e. invariante dans le temps [5]. Cependant, cela impose des contraintes importantes, nuisibles à l'expressivité. Il est plus optimal de corriger la hauteur de manière dynamique, i.e. d'adapter la fonction de correspondance entrée/sortie uniquement lorsque l'ajustement est nécessaire, autrement dit lorsque la trajectoire mélodique est stable. Cette relation entrée/sortie peut-être caractérisée par deux familles de fonctions. Les fonctions de type *notes étendues* sont les mêmes que celles utilisées pour un ajustement statique et sont largement utilisées dans des applications visuelles telles que les

“Sticky icons” [17]. Ces fonctions sont aplaties autour des notes cibles (points  $(-1, -1)$ ,  $(0, 0)$ ,  $(1, 1)$  ...) afin d'élargir les zones de justesse (figure 1 - gauche). Les fonctions de type *élastique* sont utilisées par le piano augmenté de McPherson [12] et le *Cantor Digitalis* [14]. Elles déplacent la position de la note cible sur la position d'entrée afin que toute hauteur d'entrée donne une note juste à l'activation de l'ajustement (figure 1 - droite). Contrairement à la méthode *note étendue*, une fonction *élastique* est calculée pour chaque position d'entrée. Le but de ce papier est de comparer les méthodes dynamiques *notes étendues* et *élastique*.

## 2. MÉTHODES D'AJUSTEMENT

Un ajustement dynamique est décrit en deux étapes. Il est d'abord nécessaire de définir le type de relation entre hauteurs d'entrée et de sortie (*notes étendues* vs. *élastique*). Il faut ensuite décrire le mode de déclenchement de l'ajustement, en fonction de la dynamique de la trajectoire mélodique. Le mode de déclenchement est similaire pour les deux méthodes.

### 2.1. Fonctions de correspondance statiques

Les fonctions de correspondance sont définies depuis l'espace des hauteurs d'entrée vers l'espace des hauteur de sortie. Les fonctions sont exprimées en demi-tons (ST) et relativement à la note cible courante, où le point  $(0, 0)$  correspond à la hauteur d'entrée jouant la cible exactement.  $(-1, -1)$  et  $(1, 1)$  désignent respectivement les notes précédente et suivante. Les fonctions sont tracées en figure 1.

La fonction *notes étendues* vérifie 3 conditions :

- Passe par le point  $(0, 0)$  correspondant à la cible à atteindre
- Est continûment dérivable aux points  $(-0.5, -0.5)$  et  $(0.5, 0.5)$  pour assurer une transition lisse entre les notes
- Choix du degré de distorsion de la fonction

En appelant  $z$  le coefficient de distorsion ( $z \geq 1, z \in \mathbb{R}$ ), la hauteur de sortie  $y_{NE}$  s'exprime en fonction de la hauteur

1. <http://www.rogerlinndesign.com/linnstrument.html>

2. <http://www.wizdommusic.com/products/morphwiz.html>

3. <http://www.apple.com/fr/ios/garageband/>

d'entrée  $x$  sous la forme :

$$y_{NE}(x) = 2^{z-1} \times \text{signe}(x) \times |x|^z \quad (1)$$

La fonction *élastique* vérifie 2 conditions :

- Passe par le point  $(x_0, 0)$  donnant une note juste pour la position initiale du stylet
- Passe par les points  $(-1, -1)$  et  $(1, 1)$  pour assurer la continuité avec les notes précédente et suivante

En appelant  $\gamma$  la courbure de la fonction, la hauteur de sortie  $y_E$  s'exprime en fonction de la hauteur d'entrée  $x$  sous la forme :

$$\begin{cases} y_E(x) = \frac{1}{\gamma} [\log [(e^{2\gamma} - 1)(\frac{x+1}{2}) + 1]] - 1 & \text{si } \gamma \neq 0 \\ y_E(x) = x & \text{si } \gamma = 0 \end{cases} \quad (2)$$

Pour chaque nouvel ajustement, la condition pour avoir une note juste à la hauteur initiale  $x_0$  est  $y_E(x_0, \gamma_0) = 0$ , où  $\gamma_0$  est la courbure à l'instant d'ajustement. Cela conduit à l'expression :

$$\gamma_0 = \log \left( \frac{1 - x_0}{1 + x_0} \right) \quad (3)$$

## 2.2. Dynamique de l'ajustement

L'expressivité mélodique se traduit par des variations de hauteur. A l'inverse, les notes à corriger sont supposées stables. L'ajustement est donc appliqué uniquement pour des variations faibles de hauteur, et enlevé en cas contraire. La vitesse de la hauteur d'entrée permet donc de distinguer les notes à corriger des variations expressives. On définit la vitesse moyenne critique  $AV_c$  de la hauteur comme seuil au delà duquel l'ajustement ne sera pas appliqué. Celle-ci s'exprime comme le temps critique  $T_c$  nécessaire pour parcourir un intervalle  $I$  appelé intervalle de détection :  $AV_c = I/T_c$ . Ainsi, si la hauteur d'entrée reste dans un intervalle de détection  $I$  plus longtemps que  $T_c$ , la vitesse de la hauteur mélodique est inférieure à  $AV_c$  et l'ajustement est appliqué. Inversement, si la hauteur mélodique quitte l'intervalle  $I$  avant  $T_c$ , la vitesse de la hauteur est supérieure à  $AV_c$  et l'ajustement n'est pas réalisé. L'intervalle de détection  $I$  et le temps critique  $T_c$  sont donc des paramètres pour le réglage de la dynamique d'application de l'ajustement.

L'intervalle de détection  $I$  est obtenu en divisant l'axe des hauteurs d'entrées en un nombre entier d'intervalles par demi-ton (ST). La taille de l'intervalle de détection est directement lié à la fréquence d'ajustement. La hauteur d'entrée est susceptible de changer d'intervalle de détection plus souvent lorsque ceux-ci sont petits (p. ex. 0.1 ST) que lorsque ceux-ci sont grands (0.5 ST). Un nombre plus important d'ajustements aura lieu avec de petits intervalles, favorisant la justesse. A l'inverse, de grands intervalles permettent plus de libertés de variations expressives. Cependant les variations non-intentionnelles au sein des intervalles ne sont pas corrigées.

Une fois l'intervalle de détection choisi, le temps critique permet d'ajuster le seuil de vitesse  $AV_c$ . Un faible vibrato d'amplitude  $\pm 0.1$  ST et de fréquence 5 Hz donne une vitesse moyenne  $AV = 2$  ST/s. Pour préserver les vibratos, la vitesse moyenne critique  $AV_c$  doit donc être inférieure. On peut imaginer des glissandos plus lents (p. ex. 1 ST/s). Il faut alors diminuer la vitesse moyenne critique en augmentant le temps critique pour préserver les glissandos. Plus le temps critique est long plus les variations expressives lentes sont préservées, mais moins les ajustements sont réactifs.

Pour ne pas introduire de sauts brusques dans la hauteur, l'ajustement doit s'appliquer graduellement. La fonction de correspondance est donc modifiée continûment d'une relation linéaire (sans ajustement) à la fonction voulue en un temps de transition prédéfini  $T_t$ . Cela se fait en modifiant continûment le degré de distorsion  $z$  de la fonction *notes étendues* ou la courbure  $\gamma$  de la fonction *élastique*.

Le maintien de l'ajustement après son application diffère selon la méthode. L'aplatissement de la fonction *notes étendues* autour de la cible rend les variations de hauteur fines telles que le vibrato difficiles à produire. En effet, toute oscillation de la hauteur d'entrée dans la zone aplatie résulte en une hauteur de sortie constante. C'est pourquoi l'ajustement *notes étendues* ne doit pas être maintenu durant un vibrato. Il est donc supprimé à la moindre variation de hauteur, i.e. lorsque celle-ci quitte l'intervalle de détection ayant déclenché l'ajustement. A l'inverse, la fonction *élastique* distord peu les modulations de hauteur. Elle peut donc être maintenue durant toute variation expressive. Elle est alors retirée uniquement en cas de changement de note, i.e. lorsque la hauteur quitte l'intervalle  $[-1, 1]$  autour de la note cible.

## 3. COMPARAISON

Les ajustements de justesse mélodique sont développés et testés dans le contexte du *Cantor Digitalis*<sup>4</sup>. Il s'agit d'un synthétiseur de voix chantée contrôlé en temps réel par le biais d'une tablette graphique (Wacom Intuos). Seul le contrôle de la hauteur mélodique est étudié ici. Cette dernière est modifiée continûment par la position horizontale du stylet sur la tablette. L'axe vertical est laissé libre pour contraindre au minimum le geste du musicien. Pour aider le joueur à viser les notes précisément, le support visuel montré en figure 2 est placé sur la surface de la tablette. Bien que ce dernier soit inspiré d'un clavier de piano, les "touches" sont uniquement des indicateurs visuels de la note la plus proche. L'instrumentiste doit viser les lignes pour jouer juste. Une hauteur intermédiaire sera jouée dans le cas contraire. Dans cette étude, un demi-ton correspond à 6 mm sur la tablette, et 35 demi-tons sont accessibles, de sol #

4. <http://cantordigitalis.limsi.fr>

à sol. La résolution spatiale de la position du stilet est de 0.005 mm, et fournit donc une résolution de hauteur de 0.08 centièmes de demi-tons, ce qui est grandement inférieur au seuil de perception de hauteur différentes (environ 1 centième de demi-ton à 440 Hz [13]).

Dans l'expérience suivante, le *Cantor Digitalis* est utilisé pour enregistrer les sujets, avec et sans ajustement. Les performances des deux méthodes sont comparées sur deux points : la préservation d'expressivité, et l'apport de justesse.

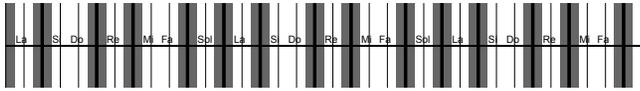


Figure 2. Calque appliqué sur la tablette

### 3.1. Expressivité

Une évaluation formelle de l'expressivité est difficile car cela implique la sensibilité musicale des sujets, ce qui est grandement subjectif. L'expressivité est donc analysée par observation de trajectoires contenant les trois principaux effets de modulation de hauteur : *portamento*, *glissando* et *vibrato*. Idéalement, un ajustement de justesse doit à la fois donner une hauteur juste et précise mais aussi préserver les trajectoires des modulations. Trois réglages d'ajustement sont étudiés et présentés en table 1.

Table 1. Paramètres de déclenchement utilisés pour l'étude de l'expressivité

	Intervalle de de détection $I$	Temps critique $T_c$	Vitesse moyenne critique $AV_c$
<b>A</b>	0.1 ST	0.1 s	1 ST/s
<b>B</b>	0.5 ST	0.25 s	2 ST/s
<b>C</b>	0.1 ST	0.25 s	0.4 ST/s

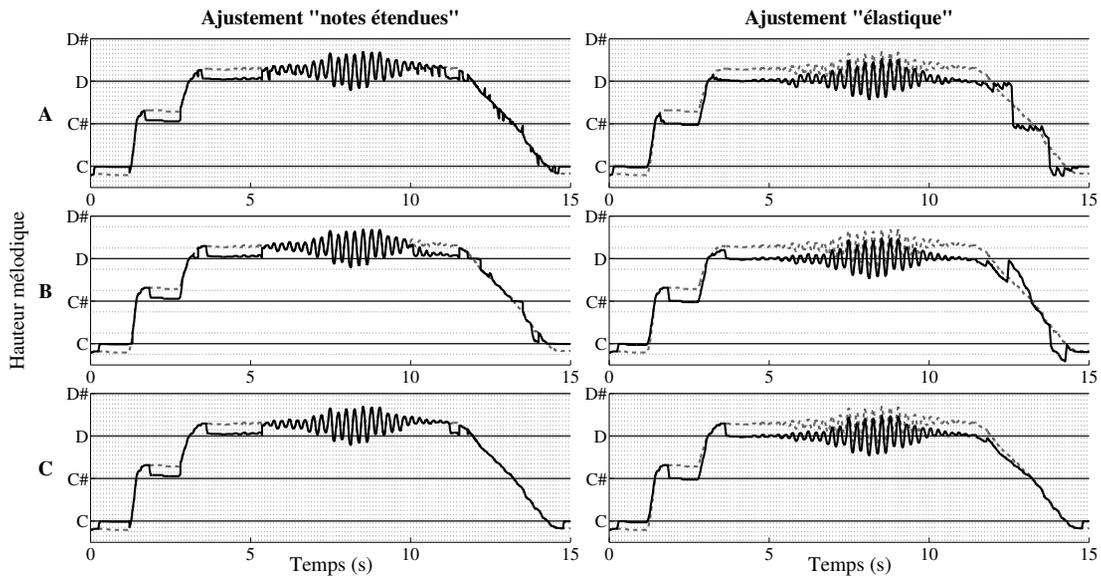
Le réglage A est un ajustement "rapide" caractérisé par un petit intervalle de détection et un faible temps critique. A l'inverse le réglage B est un ajustement "lent". Son intervalle de détection plus grand et son long temps critique permettent plus de liberté de variations expressives. Le réglage C possède un petit intervalle de détection et un long temps critique. Un temps de transition  $T_t = 50$  ms est utilisé pour les 3 réglages. Ces ajustements sont appliqués à une même mélodie reflétant les trois modulations détaillées plus haut (portamento, glissando et vibrato). Les étapes de la mélodie sont : début sur un Do (C) ; deux portamentos successifs vers les notes Do # (C#) et Ré (D) ; vibrato ; glissando descendant vers la note Do (C). Les valeurs d'entrées et de sorties sont représentées en figure 3, en courbe pointillée et en courbe pleine respectivement. Les lignes pleines horizontales sont les notes cibles, les lignes pointillées horizontales sont les limites des intervalles de détection.

Une première observation montre que le portamento est bien ajusté pour chaque réglage, et chaque méthode d'ajustement. Seul le temps de réaction diffère, directement lié au temps critique ( $T_c = 100$  ms pour le réglage A, et  $T_c = 250$  ms pour les réglages B et C). En revanche, le vibrato est centré autour de la note cible uniquement par l'ajustement *élastique*. Cela est lié aux conditions de suppression de l'ajustement pour les deux méthodes (section 2.2). Dans le cas de la méthode *notes étendues*, l'ajustement est supprimé dès que la hauteur quitte l'intervalle de détection ayant déclenché l'ajustement. C'est pourquoi dans le cas de vibratos d'amplitudes supérieures à l'intervalle de détection, l'ajustement n'est pas conservé.

En haut figurent les résultats de l'ajustement rapide (réglage A). Ces faibles valeurs induisent un déclenchement fréquent de l'ajustement, principalement pendant les glissandos. A chaque changement d'intervalle de détection, la hauteur est ajustée en introduisant des distorsions. Dans le cas de la méthode *notes étendues* l'ajustement est successivement appliqué, puis retiré, et ainsi de suite, créant des pics dans la trajectoire. Dans le cas de la méthode *élastique*, l'ajustement est maintenu toute la durée du glissando, conduisant à l'apparition de palliers. Concernant le vibrato, l'ajustement est bien supprimé pour la méthode *notes étendues*, à 5.4 s. De plus, on observe de très légères distorsions au début et à la fin des vibratos pour chaque méthode. Cela ce produit lorsque l'amplitude du vibrato devient proche de la taille de l'intervalle de détection. Comme le temps critique  $T_c = 100$  ms est inférieur à une période de vibrato (150-200 ms), la hauteur d'entrée reste suffisamment longtemps dans un intervalle de détection pour déclencher l'ajustement au sein même d'un cycle de vibrato. Ces réglages sont donc à la limite de préservation des vibratos. Un temps critique plus court entraînerait des plus grandes distorsions. Finalement, si les ajustements sont trop rapides, ils deviennent intrusifs et empêchent certaines formes d'expressivité.

Les résultats de l'ajustement lent (réglage B) sont montrés au milieu de la figure 3. De même qu'avec le réglage A, l'ajustement *notes étendues* est à nouveau supprimé à 5.4 s pour le vibrato. Il est cependant intéressant de noter qu'après 10 s, le vibrato est inclus dans un intervalle de détection, et est donc ajusté. Des intervalles de détection plus larges sont donc nécessaires pour ajuster les vibratos avec la méthode *notes élargies*, au dépend de distorsions introduites par la fonction de correspondance. Les paramètres d'ajustement induisent une vitesse moyenne critique de 2 ST/s, supérieure à la vitesse du glissando. Cela explique donc les distorsions observées.

Le réglage C est défini par le petit intervalle de détection du réglage A ( $I = 0.1$  ST) et le long temps critique du réglage B ( $T_c = 250$  ms). Le bas de la figure 3 montre les résultats des deux méthodes avec ce réglage. Aucune distorsion n'apparaît pour chacun des effets avec les deux mé-



**Figure 3.** Exemples d’ajustements (gauche : *notes étendues* ; droite : *élastique*) avec différents réglages (voir table 1). Les courbes en pointillés sont les trajectoires de la hauteur d’entrée. Les courbes pleines sont les trajectoires des hauteurs de sortie. Les lignes pleines sont les notes cibles. Les lignes pointillées représentent les intervalles de détection  $I$  (A et C : 0.1 ST, B : 0.5 ST).

thodes. Comme précédemment, le vibrato n’est pas ajusté avec la méthode *notes étendues*, mais ce dernier n’est pas distordu.

La comparaison entre les réglages B et C indique que dans le cas de l’ajustement *élastique*, un petit intervalle de détection entraîne une meilleure détection des zones stables de la trajectoire mélodique. Cela permet à la fois un ajustement fréquent de la trajectoire et une préservation de toutes les formes expressives. A l’inverse, dans le cas de l’ajustement *notes étendues*, un grand intervalle de détection est nécessaire pour l’ajustement du vibrato, au dépend de la fréquence d’ajustement.

La comparaison entre les réglages A et C montre qu’un temps critique long permet de mieux préserver les modulations expressives. Néanmoins, l’enchaînement des notes est lent. Pour des tempos plus rapides, en temps critique plus court est nécessaire, au dépend des modulations expressives (réglages A).

Pour conclure, les deux méthodes introduisent un minimum de distorsion sur les trajectoires mélodiques tout en corrigeant la justesse efficacement, sous réserve que des paramètres appropriés soient choisis (réglage C). Cependant, la méthode *notes étendues* introduit un compromis entre fréquence d’ajustement et correction et distorsion des vibratos. Elle est donc moins robuste que la méthode *élastique* pour la préservation de l’expressivité.

## 3.2. Justesse et précision

### 3.2.1. Protocole

La justesse apportée par les ajustements est quantifiée dans un deuxième temps par expérience d’imitation sur des tâches techniques. Les sujets doivent reproduire des mélodies à l’aide d’un stylet sur une tablette graphique. Quatre mélodies sont proposées dans différentes conditions : sans ajustement, avec l’ajustement *notes étendues* et avec l’ajustement *élastique*. Pour chaque stimulus, la mélodie est jouée par un synthétiseur MIDI (son de piano, A4 = 440 Hz) autant de fois que nécessaire, et une partition contenant notation musicale et nom des notes est fournie. Les stimuli sont présentés aléatoirement à travers un casque audio à un volume confortable. Chaque stimulus est joué 3 fois par chaque sujet. Il est demandé aux sujets de ne pas rompre le contact entre stylet et surface de la tablette pendant la durée d’un essai afin de tester les différentes dynamiques d’ajustement. Pour chaque correction, deux réglages sont testés : un ajustement rapide ( $I = 0.1$  ST ;  $T_c = 0.1$  s ;  $T_t = 0.05$  ms) et un ajustement lent ( $I = 0.5$  ST ;  $T_c = 0.25$  s ;  $T_t = 0.05$  ms). Ils correspondent aux réglages A et B de la table 1. Ces quatre conditions ainsi qu’une condition “sans ajustement” sont proposées pour chaque motif mélodique. Celles-ci ne sont pas communiquées aux sujets. Les mélodies sont des motifs de 9 notes inspirés d’exercices vocaux et sont présentées en figure 4. Deux tempos différents sont utilisés : 120 et 240 battements par minute (b.p.m.) et indiqués par un métronome de manière visuelle et auditive. Au total, chaque

sujet doit jouer 3 fois 4 mélodies sous 5 conditions d'ajustement pour 2 tempos.



Figure 4. Mélodies utilisées dans l'expérience

10 sujets (29 ans et 12 années d'expérience musicale en moyenne) ont participé à l'expérience. Aucun des sujets ne présente de déficience auditive et tous sont droitiers. Une tablette Wacom 5M équipée du calque présenté en figure 2 a été utilisée pour l'expérience. Chaque sujet a été informé des modalités de l'expérience et a réalisé une session d'entraînement au préalable, présentant les mêmes stimuli et protocole que l'expérience.

### 3.2.2. Analyse

Pour chaque essai, les notes jouées par les sujets ont été identifiées à la fois sur les hauteurs d'entrée et de sortie comme les parties stables entre les transitions inter-notes (pics dans la dérivée). Pour chaque note, deux valeurs sont extraites :

- La valeur d'entrée : extraite de la hauteur d'entrée avant l'application de l'ajustement (la note jouée lorsque le stylet est stable).
- La valeur de sortie : extraite de la hauteur de sortie après une éventuelle application de l'ajustement (la note entendue par le sujet).

Dans le cas où l'ajustement est attendu mais pas réalisé, les notes concernées sont retirées et le pourcentage des notes restantes est discuté plus bas.

L'analyse de la performance des ajustements repose sur les notions de justesse et précision. Pour un ensemble de notes données, on calcule d'abord l'erreur associée à chaque note comme la différence entre la note jouée et la note à atteindre correspondante. La justesse est la valeur moyenne de ces erreurs. Une valeur de justesse nulle indique que les erreurs du musicien se compensent. Une valeur de justesse positive indique un jeu globalement trop haut et inversement. La précision est l'écart-type des erreurs. Une valeur de précision nulle indique une erreur constante pour chaque note. Une valeur de précision positive indique des erreurs incon-

sistantes selon les notes. Justesse et précision ont été proposées et utilisées pour l'évaluation de performances chantées [15]. Chaque sujet a joué l'ensemble des stimuli 3 fois. Par conséquent, chaque réalisation de l'ensemble des stimuli sera considérée séparément pour donner un couple de valeurs justesse précision par sujet et par réalisation.

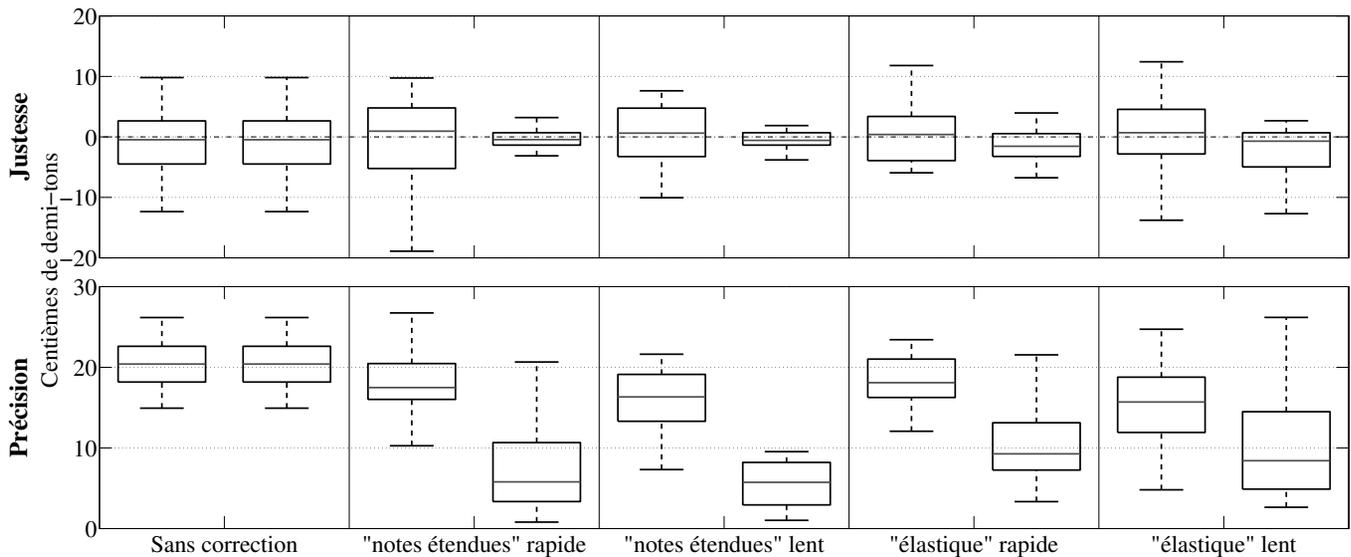
### 3.2.3. Résultats

La figure 5 montre la justesse (haut) et la précision (bas) exprimées en centièmes de demi-tons (cents) des sujets dans les 5 conditions d'ajustement. Pour chaque condition, la boîte de gauche contient les justesses (resp. précisions) des valeurs d'entrée de chaque réalisation de chaque sujet, et la boîte de droite contient les justesses (resp. précisions) des valeurs de sortie de chaque réalisation de chaque sujet. Les différences statistiques entre justesse et précision parmi les différents ajustements sont étudiées avec un test de Wilcoxon par paires.

Les justesses obtenues pour chaque condition ont toutes une médiane proche de 0 et une dispersion inférieure à 5 cents. Ces résultats concordent avec l'étude du contrôle de la voix chantée [4]. Les valeurs d'entrées des joueurs étant déjà très justes, la justesse n'est pas significativement améliorée pour aucun des ajustements. Les médianes des précisions des valeurs d'entrée sont situées entre 15 et 20 cents. Cela correspond à une distance d'environ 1 mm sur la tablette, de l'ordre de grandeur de la largeur de la pointe du stylet. On a donc une précision proche de la limite imposée par la tâche. Toutes les précisions des valeurs de sortie sont en dessous de 15 cents et sont significativement plus faibles pour chacun des ajustements : *notes étendues* rapide ( $W = 845, p < 0.001$ ) ; *notes étendues* lent ( $W = 761, p < 0.001$ ) ; *élastique* rapide ( $W = 804, p < 0.001$ ) ; *élastique* lent ( $W = 664, p < 0.01$ ).

On observe ensuite un effet de l'ajustement sur les précisions des valeurs de sortie. La valeur de précision fournie par l'ajustement *notes étendues* rapide (resp. lent) et significativement plus faible que la valeur de précision fournie par l'ajustement *élastique* rapide ( $W = 242, p < 0.01$ ) (resp. lent ( $W = 290, p < 0.05$ )). L'ajustement *notes étendues* définit une zone autour de la note cible dans laquelle toutes les positions d'entrées produisent cette cible. L'ajustement *élastique* déplace la cible vers la position d'entrée actuelle, mais autorise de faibles déviations autour de la position corrigée. Par conséquent ce dernier est plus sensible aux faibles mouvements et conduit à une correction moins stable.

Enfin, aucune différence significative n'apparaît entre ajustements lents et rapides. Cependant le nombre de notes corrigées avec les ajustements rapides (57%) est supérieur au nombre de notes corrigées avec les ajustements lents (38%). Les ajustements rapides ont un temps critique de 100 ms et un temps de transition de 50 ms. Ils sont donc effectifs après 150 ms. Les ajustements lents sont effectifs après 250



**Figure 5.** Justesse (haut) et précision (bas) des sujets en fonction de la condition d’ajustement. Pour chaque condition la boîte de gauche (resp. droite) contient les valeurs d’entrée (resp. sortie) de chaque réalisation.

ms. Un tempo de 240 b.p.m. donne une pulsation toutes les 250 ms. Cela laisse donc peu de temps pour les ajustements après stabilisation de la note. Par conséquent, les paramètres d’ajustement définissent un tempo limite au-dessus duquel les ajustements ne sont plus efficaces.

Pour conclure, les deux méthodes d’ajustement améliorent significativement la précision des joueurs, avec de meilleures performances pour l’ajustement *notes étendues*. Néanmoins, la précision des valeurs de sortie est proche de la limite de perception de la hauteur dans tous les cas. Les deux méthodes sont toutes deux limitées par leurs temps de réaction, qui imposent une longueur de note suffisante pour que celles-ci soient efficaces.

#### 4. DISCUSSION ET CONCLUSION

Deux méthodes d’ajustement de justesse dynamiques redéfinissant la relation entre hauteurs d’entrée et de sortie ont été développées et comparées. Du point de vue la justesse, il a été montré que les deux méthodes permettent d’atteindre des justesses et précisions proches ou en dessous du seuil de perception, et inférieures à la limite imposée par la largeur de la pointe du stylet. Quant aux modulations expressives, le choix des paramètres s’avère primordial. Pour un ajustement réactif, nous suggérons l’utilisation d’un intervalle de détection  $I = 0.1$  ST. Le temps critique doit être ajusté en fonction du tempo de la pièce. La section 3.1 montre que  $T_c = 250$  ms entraîne le moins de distorsions pour un tempo lent. Un temps critique plus petit doit être choisi pour des tempos plus grands, comme il a été montré en section 3.2. La limitation de l’expressivité introduite par un temps cri-

tique court est compensée par le fait qu’un musicien montre naturellement moins d’expressivité mélodique lors de successions de notes rapides tout en conservant d’autres formes d’expressivité.

Malgré le choix de réglages optimaux, seule la méthode *élastique* permet d’éviter toute distorsion pour les trois types de modulations les plus courantes. On en conclut donc que la fonction de correspondance *élastique* entre hauteurs d’entrée et de sortie est plus favorable à l’expressivité mélodique. Comme sur un piano, la hauteur est juste pour n’importe quelle position du stylet au moment de l’ajustement. Une fois la note corrigée, des modulations libres sont possibles comme sur un violon.

#### 5. REFERENCES

- [1] d’Alessandro, N. and Dutoit, T., « Handsketch : Bi-Manual Control of Voice Quality Dimensions and Long Term Practice Issues », *Numediart Research Program*, Mons, Belgium, 2009.
- [2] d’Alessandro, N. and Dutoit, T., « Advanced Techniques for Vertical Tablet Playing A Overview of Two Years of Practicing the HandSketch 1.x », *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Pittsburg, PA, USA, 2009, p. 173–174.
- [3] d’Alessandro, C. and Rilliard, A. and Le Beux, S., « Chironomic Stylization of Intonation », *Acoustical Society of America*, 2011, p. 1594–1604.
- [4] d’Alessandro, C. and Feugère, L. and Le Beux, S. and Perrotin, O. and Rilliard, A., « Drawing melodies : Eva-

- luation of Chironomic Singing Synthesis », *Acoustical Society of America*, 2014, p. 3601–3612.
- [5] Goudard, V. and Genevois, H. and Feugère, L., « On the playing of monodic pitch in digital music instruments », *ICMC/SMC 2014*, Athènes, Greece, 2014, p. 1418–1425.
- [6] Haken, L. and Tellman, E. and Wolfe, P., « An Indiscrete Music Keyboard », *Computer Music Journal*, 1998, p. 30–48.
- [7] Haken, L., « Position correction for an electronic musical instrument », *US Patent 7,619,156*, 2009.
- [8] Juslin, P. N. and Laukka, P. « Communication of Emotions in Vocal Expression and Music Performance : Different Channels, Same Code ? », *Psychological Bulletin*, 2003, p. 770–814.
- [9] Kessous, L., « Gestural Control of Singing Voice, a Musical Instrument », *Proceedings of Sound and Music Computing*, Paris, France, 2004.
- [10] Lamb, Roland and Robertson, Andrew N., « Seabord : a New Piano Keyboard-Related Interface Combining Discrete and Continuous Control », *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Oslo, Norway, 2011, p. 503–506.
- [11] Le Beux, S. and d’Alessandro, C. and Rilliard, A. and Doval, B., « Calliphony : A System for Real-time Gestural Modification of Intonation and Rhythm », *Speech Prosody*, Chicago, IL, USA, 2004.
- [12] McPherson, A. P. and Gierakowski, A. and Stark, A. M., « The Space Between the Notes : Adding Expressive Pitch Control to the Piano Keyboard », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2013, p. 2195–2204.
- [13] Moore, B. C. J., « Frequency difference limens for short-duration tones », *Acoustical Society of America*, 1973, p. 610–619.
- [14] Perrotin, O. and d’Alessandro, C., « Adaptive mapping for improved pitch accuracy on touch user interfaces », *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Daejeon, South Korea, 2013, p. 186–189.
- [15] Pfordresher, P. Q. and Brown, S. and Meier, K. M. and Belyk, M. and Liotti, M., « Imprecise singing is widespread », *Acoustical Society of America*, 2010, p. 2182–2190.
- [16] Wanderley, M. M. and Violette, J.-P. and Isart, F. and Rodet, X., « On the Choice of Transducer Technologies for Specific Musical Functions », *Proceedings of the International Computer Music Conference (ICMC)*, Ann Arbor, MI, USA, 2000.
- [17] Worden, A. and Walker, N. and Bharat, K. and Hudson, S., « Making Computers Easier for Older Adults to Use : Area Cursors and Sticky Icons », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2000, p. 266–271.
- [18] Zbyszynski, M. and Wright, M. and Momeni, A. and Cullen, D., « Ten Years of Tablet Musical Interfaces at CNMAT », *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, New York, NY, USA, 2007, p. 100–105.