

L'IRCAM et la voix augmentée au théâtre : les nouvelles technologies sonores au service de la dramaturgie

Grégory Beller

Numéro 56-57, automne 2014, printemps 2015

Écouter la scène contemporaine

URI : <https://id.erudit.org/iderudit/1037338ar>

DOI : <https://doi.org/10.7202/1037338ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société québécoise d'études théâtrales (SQET)
Université de Montréal

ISSN

0827-0198 (imprimé)
1923-0893 (numérique)

[Découvrir la revue](#)

Citer cet article

Beller, G. (2014). L'IRCAM et la voix augmentée au théâtre : les nouvelles technologies sonores au service de la dramaturgie. *L'Annuaire théâtral*, (56-57), 195–205. <https://doi.org/10.7202/1037338ar>

Résumé de l'article

Cet article vise à rendre publiques de nouvelles solutions créatives pour la dramaturgie sonore dans le spectacle vivant. Ces nouvelles technologies sont le fruit d'une trentaine d'années de recherches sur la voix menées par plusieurs équipes de l'Institut de recherche et coordination acoustique / musique (IRCAM). Dans un premier temps, la problématique générale du traitement automatique de la voix est abordée afin de situer la spécificité de l'IRCAM dans le contexte international de la recherche sur la parole et dans le but de présenter différents facteurs qui rendent, aujourd'hui, les nouvelles technologies sonores bénéfiques à la création artistique. Dans un deuxième temps, cet article tente de présenter différents outils qui peuvent apporter des éléments de réponse à des questions dramaturgiques. Enfin, la dernière partie de l'article porte une attention particulière au suivi de voix, jeune technologie prometteuse qui sert la composition théâtrale et la régie sonore. Cet article se termine par un rappel de l'importance d'intégrer ces outils dans le processus de création, les distinguant par là des techniques d'amplification palliative du théâtre.

GRÉGORY BELLER

Artiste, chercheur et
directeur des interfaces
recherche-cr  ation de l'IRCAM

L'IRCAM et la voix augment  e au th   tre : les nouvelles technologies sonores au service de la dramaturgie

Le th   tre a toujours   t   le lieu de l'innovation technique, notamment dans le domaine du son. De l'architecture des amphith   tres grecs en passant par les machines    effet sonore du XIX   si  cle jusqu'au th   tre radiophonique, de nombreux auteurs et metteurs en sc  ne ont syst  matiquement exploit   les nouvelles connaissances de leur   poque. En t  moigne, aujourd'hui, le large spectre de cr  ateurs utilisant les techniques modernes de captation, de traitement et de diffusion du son : Robert Wilson, Romeo Castellucci, Christoph Marthaler, Heiner Goebbels, Jean-Fran  ois Peyret, etc.

Depuis sa cr  ation, l'Institut de recherche et coordination acoustique / musique (IRCAM) a pour vocation de frotter les artistes aux chercheurs scientifiques dans le but de permettre    chacun de nourrir sa pratique du fantasme de l'autre. Jusqu'ici, et par tradition, la plupart des artistes actifs de l'IRCAM ont   t   des compositeurs de musique mixte¹. Ensemble, avec les scientifiques, ils ont identifi   des probl  matiques r  currentes de la composition sonore et ont pouss   l'innovation technique jusqu'   produire de v  ritables applications g  n  ralistes. Ces applications d  passent d  sormais le seul cadre de la musique mixte puisqu'elles sont employ  es dans la postproduction cin  matographique ou dans le design sonore en g  n  ral (Schwarz, Beller, Verbrugge et Britton, 2006).

1 La musique mixte allie des sons   manant d'instruments de musique – des sons instrumentaux –    des sons   lectroacoustiques. (Note de la r  daction.)

La rencontre entre l'IRCAM et le spectacle vivant était donc inévitable, voire prédestinée. Aussi, de plus en plus de metteurs en scène viennent participer aux activités de l'Institut : Joël Pommerat, Stanislas Nordey, Ludovic Lagarde, Georges Aperghis, François Verret, Matthieu Roy, Cyril Teste, Émilie Rousset ou Guillaume Vincent par exemple. Cela, notamment, grâce à la politique d'ouverture de l'IRCAM qui se traduit par des coproductions, un forum d'utilisateurs et des laboratoires d'expérimentation autour du son dans le spectacle vivant appelés In Vivo.

Cet article vise à rendre publiques de nouvelles solutions créatives pour la dramaturgie sonore dans le spectacle vivant. Ces nouvelles technologies sont le fruit d'une trentaine d'années de recherches sur la voix menées par plusieurs équipes de l'IRCAM. Dans un premier temps, la problématique générale du traitement automatique de la voix est abordée afin de situer la spécificité de l'IRCAM dans le contexte international de la recherche sur la parole et dans le but de présenter différents facteurs qui rendent, aujourd'hui, les nouvelles technologies sonores bénéfiques à la création artistique. Dans un deuxième temps, à partir d'exemples concrets tirés de spectacles, cet article tente de présenter différents outils qui peuvent apporter des éléments de réponse à des questions dramaturgiques. Une troisième partie accorde une attention particulière au suivi de voix, jeune technologie prometteuse pour la composition théâtrale et la régie sonore. Cet article se conclut par un rappel de l'importance d'intégrer ces outils dans le processus de création, les distinguant par là des techniques d'amplification palliative du théâtre.

L'IRCAM ET LA VOIX

De nombreuses technologies sonores issues de projets de recherche et développées à l'IRCAM peuvent servir à la création sonore pour le théâtre, comme la spatialisation sonore par exemple. Dans cet article, nous ne dresserons pas une liste exhaustive de tous ces outils. Nous nous concentrerons plutôt sur les outils relatifs à la voix, source sonore primordiale au théâtre. La voix constitue l'objet sonore le plus complexe qui soit, tant au niveau de sa variabilité acoustique que dans son rapport au sens, à l'intimité et à l'individualité du locuteur. De plus, nous sommes tous des experts de la voix puisque c'est le type de source sonore que nous entendons le plus au quotidien. Ce constat oblige quiconque veut traiter la voix par la machine à le faire de manière très subtile et nuancée ainsi qu'à minimiser tout artefact immédiatement perçu par un auditeur lambda. Afin d'expliquer la ligne de conduite des développements effectués par les chercheurs de l'IRCAM, prenons l'exemple de la synthèse vocale qui, de tous les domaines de recherche sur la parole, est certainement celui qui a permis le plus d'innovations techniques ces dernières années.

En effet, depuis les années 1950, l'informatique a permis l'accélération du traitement automatique de la parole et du langage. De nombreux laboratoires ont effectué des recherches pour analyser, reconnaître, transformer ou synthétiser le signal vocal. De nombreuses applications ont découlé de ces recherches permettant, pour la plupart, la translation entre le son et le sens. Ainsi, les techniques de reconnaissance vocale et de synthèse à partir du texte réalisent aujourd'hui une interface plus ou moins naturelle entre l'homme et la machine.

Ces techniques font souvent appel à de grands corpus de parole annotée. Depuis une quinzaine d'années environ, l'IRCAM a développé de tels corpus (Veaux, Beller et Rodet, 2008b) parallèlement aux outils permettant de les exploiter (Beller, Veaux, Degottex, Obin, Lanchantin et Rodet, 2008; Veaux, Beller et Rodet, 2008a). En synthèse de la parole, la spécificité des recherches de l'IRCAM par rapport à celles conduites par d'autres laboratoires repose sur la qualité sonore de la voix produite. Là où certains ont concentré leurs efforts sur la compression des voix pour des échanges téléphoniques ou la lecture rapide d'Internet, les équipes de l'IRCAM ont cherché à produire des voix d'une qualité irréprochable pour une utilisation en contexte artistique. De plus, elles ont créé des outils permettant un contrôle musical de la prosodie de la voix de synthèse (Beller, 2005; Beller, Schwarz, Hueber et Rodet, 2005). Cette façon unique et singulière d'élaborer un synthétiseur de parole a permis à ce dernier d'être utilisé dans le spectacle *Luna Park*² de Georges Aperghis. La synthèse vocale composée y remplit une fonction de voix off, allégorie d'une voix toute-puissante de la machine (allusion à *Big Brother*). L'utilisation de la voix de synthèse peut donc remplir des fonctions didascaliques ou narratives en voix off tout en jouant un rôle particulier de substitution du caractère humain, alors qu'elle utilise paradoxalement ce qui, peut-être, en fait sa quintessence, la voix.

D'autres outils reposent sur une quasi-instantanéité désormais possible du traitement de la voix. Les applications « temps réel » développées à l'IRCAM permettent d'analyser, de suivre, de transformer et de démultiplier non seulement la voix chantée, mais aussi la parole et, plus largement, tout geste vocal. Offrant un délai de traitement inférieur à notre seuil de perception de la discontinuité sonore (de l'ordre de 20ms), ces techniques sont aujourd'hui utilisées dans le cadre du spectacle vivant impliquant la voix amplifiée. Elles permettent à la fois aux interprètes d'« augmenter » leur voix et aux metteurs en scène d'apporter de nouveaux éléments de réponse à des questions dramaturgiques.

2 *Luna Park* a été créé en 2011 au Festival ManiFeste dans une mise en scène de Georges Aperghis (musique de Georges Aperghis; mise en son de Grégory Beller).

LA VOIX AUGMENTÉE POUR LE THÉÂTRE

La facture nouvelle des instruments augmentés provient des traditions croisées de la musique contemporaine mixte et de la musique électronique improvisée. Le son d'un instrument, c'est-à-dire le son d'une source sonore physique jouée par un instrumentiste, est complété / augmenté par une source sonore électronique dont l'évolution est relative au jeu de l'instrumentiste. Par exemple, dans le cas d'un violon augmenté, on peut choisir d'ajouter une réverbération au son, dont les paramètres vont dépendre de la hauteur des notes jouées. On peut aussi choisir d'ajouter un capteur de mouvements, situé au niveau de l'archet, afin d'obtenir des informations sur le jeu de l'instrumentiste qui ne sont pas directement déductibles du son produit par le violon. Quoi qu'il en soit, la motivation centrale de la facture des instruments augmentés réside dans l'instauration d'une boucle de rétroaction entre l'instrumentiste et la machine. La machine « écoute » l'instrumentiste pour agir en conséquence, et vice versa.

En ce qui concerne la voix augmentée, il est pour le moment difficile de placer des capteurs dans l'appareil vocal de manière non invasive ou non obstructive³. Or, on peut, en revanche, grâce au microphone, capter la voix au plus près, puis, grâce à la rapidité des ordinateurs, l'analyser et la traiter en temps réel. Enfin, grâce aux systèmes de diffusion, on peut produire des sonorités qui augmentent, complètent ou remplacent même la voix directe du comédien. Ce nouveau lien de subordination de la machine à l'interprète permet à celui-ci de ré-explore son instrument vocal, tout en jouissant des nouvelles possibilités de l'augmentation. En pratique, une « bonne voix augmentée » n'est possible que si le comédien, le metteur en son et la machine travaillent en équipe afin d'inscrire la voix résultante dans une dramaturgie sonore qui fait sens.

ANALYSE DE LA PAROLE

Afin de créer une augmentation sensible au jeu du comédien, il convient de conférer à la machine de bonnes oreilles. De nombreux efforts ont été consacrés à l'analyse prosodique en temps réel. Du point de vue du chercheur, on décrit la prosodie, ou la manière de parler, selon cinq dimensions (Beller, 2009a) : l'intonation, le débit de parole, l'intensité, le degré d'articulation et la qualité vocale ou l'effort vocal. Une transcription instantanée de la prosodie en notation musicale est aujourd'hui possible si l'on considère que chaque syllabe est une note (Beller, 2009b). On peut aussi transcrire le timbre de la voix en agrégats de notes ou d'accords. Afin de permettre une telle transcription, des recherches (Beller, 2007; Beller et Rodet, 2007) ont été menées pour

3 D'autres outils non obstructifs ont été développés à l'IRCAM comme la captation du geste et du mouvement. Le contrôle gestuel de la synthèse vocale dans *Luna Park* et dans *Babil-on* en est un exemple. Voir Beller et Aperghis, 2011.

segmenter, en temps réel, le flux de la parole en syllabes, mais aussi en unités plus grandes (groupes prosodiques de souffle, de phrase ou d'activité de la parole [silence/parole]) et en unités plus petites grâce à la discrimination voisé / non voisé, qui permet une segmentation entre les voyelles et les consonnes. Ces différentes variables d'analyse de la parole en temps réel constituent la façon dont la machine écoute l'interprète. Elles vont permettre à celui-ci de jouer dynamiquement et de manière contrôlée de différents effets sonores.

Dans la seconde partie de *Nous les vagues*⁴, deux comédiens représentent des leaders politiques et parlent à une foule survoltée dont la présence se manifeste par une bande sonore. Une interaction sonore, représentée métaphoriquement par un mouvement de vagues, se produit alors entre cette foule et ces orateurs. Lorsque ces derniers parlent, la foule se tait alors qu'elle devient assourdissante lorsque les orateurs se taisent. Cette interaction a été effectuée grâce à une compression *side-chain* inversée de la bande en fonction de l'intensité vocale des comédiens. Rapidement, le combat entre les orateurs et la foule assourdissante conduit paradoxalement le spectateur à rechercher le sens dans la mêlée et à écouter le discours politique. Ce même effet a été utilisé dans *La mort de Danton*⁵.

TRANSFORMATION DE LA VOIX

Dans *Un mage en été*⁶, la voix du comédien, Laurent Poitrenaux, est constamment modulée par la machine, permettant à celui-ci de se draper de différents personnages (Beller, 2011). La transformation de l'identité vocale du locuteur devient alors un déguisement sonore au service de la dramaturgie, au même titre qu'un masque permet à un personnage d'en jouer un autre. Une succession de réglages et d'ajustements a été réalisée par le comédien et le metteur en son de manière à optimiser l'adéquation entre la voix et son augmentation, selon les personnages désirés. Parmi les possibles caractéristiques de la voix qu'est capable de modifier la machine de manière réaliste, on recense :

- L'identité du locuteur : genre, âge, état de santé, taille, caractéristiques vocales (voix tremblée, tendue, etc.);
- L'identité spécifique : transformer la voix de A spécifiquement en la voix de B;

4 *Nous les vagues* est un texte de Mariette Navaro qui a été mis en scène par Matthieu Roy en 2012 (création In Vivo; mise en son de Baptiste Poulain et de Grégory Beller).

5 *La mort de Danton* est un texte de Georg Büchner qui a été mis en scène par Ludovic Lagarde à la Comédie de Reims en 2012 (mise en son de David Bichindaritz et de Grégory Beller).

6 *Un mage en été* est un texte d'Olivier Cadiot qui a été mis en scène par Ludovic Lagarde au Festival d'Avignon en 2012 (mise en son de David Bichindaritz et de Grégory Beller).

- L'expressivité : modifications prosodiques (ambitus intonatif, débit de parole, intensité, degré d'articulation, effort vocal);
- Le nombre de locuteurs : chœur pour voix chantée, foule pour voix parlée;
- L'hybridation de la voix : « animalisation » de la voix, fusion de deux voix, substitution de la source.

Dans *Mars-Watchers*⁷, une seule comédienne, Perle Palombe, joue plusieurs personnages, soit des cosmonautes masculins qui possèdent tous une voix différente. À l'inverse, Laurent Poitrenaux, dans *Un mage en été*, prend alternativement la voix d'une femme pour faire dialoguer un personnage et sa mère.

On peut ajouter à cette liste les modifications relatives au positionnement de la source sonore dans l'espace : changement de l'espace physique (diffusion / spatialisation), variation de la localisation de la source sonore ou simulacre de variation temporelle par un vieillissement virtuel du système de diffusion / captation, comme c'est le cas dans *Un mage en été*. Dans *Bedroom Eyes*⁸, le collectif MxM propose au spectateur la vision d'un comédien sur le plateau simultanément au point de vue subjectif de celui-ci, et ce, grâce à un dispositif vidéo de réalité augmentée. L'enjeu technique pour ce projet a reposé sur la capacité du son à entraîner le spectateur dans la réalité subjective du comédien. Ainsi, c'est toute la scène sonore qui tourne autour du public lorsque le comédien tourne la tête.

Enfin, on peut ajouter tout autre effet sonore musical qui retire le caractère réaliste à la voix produite : flanger, chorus, phaser, modulation en anneaux, vocodeur de phase, granulation et *sampling*. Ce dernier est largement utilisé pour créer des effets de mémoire (à court ou à long terme) au sein de la dramaturgie sonore. Dans *La nuit tombe*⁹, un effet de granulation sur les sifflantes de la comédienne Susann Vogel permet d'accentuer l'angoisse discernable dans son jeu et de faire basculer une situation, du réel au fantasme. Dans *Mars-Watchers*, un duo musico-narratif s'entremêle grâce à l'utilisation d'un vocodeur de phase mélangeant la voix du didascale et le jeu d'un musicien, tous deux présents sur scène.

7 *Mars-Watchers* est un texte d'Anne Kawala qui a été mis en scène par Émilie Rousset en 2012 (création In Vivo; mise en son de Romain Vuillet et de Thomas Goepfer).

8 *Bedroom Eyes* est un texte de Frédéric Vossier qui a été mis en scène par Cyril Teste en 2012 (création In Vivo; mise en son de Nihil Bordures et de Thomas Goepfer).

9 *La nuit tombe* est un texte de Guillaume Vincent mis en scène par l'auteur en 2012 (création In Vivo; mise en son de Géraldine Foucauld et d'Olivier Pasquet).

DÉMULTIPLICATION DE LA VOIX

Le chœur a toujours joué une fonction dramaturgique primordiale au théâtre, voire originelle. En 2013, Joël Pommerat vient à l'IRCAM et s'empare d'un outil permettant la démultiplication d'une voix en de multiples identités pour la création d'*Une année sans été*¹⁰. Initialement conçu pour l'harmonisation d'un chœur virtuel pour voix chantée solo, cet outil permet aussi de générer une foule d'individualités proférant le même message que son utilisateur, de manière synchrone ou asynchrone. Ainsi, cet outil permet à un comédien d'incarner simultanément le coryphée et ses choreutes, provoquant de complexes faisceaux dramaturgiques relatifs aux rapports entre l'individu et le groupe. La première partie de *Nous les vagues* présente la constitution d'un tel groupe : salle au noir, deux comédiens, Philippe Canales et Johanna Silberstein, sont derrière le rideau. Leurs voix sont transformées, en temps réel, en de multiples personnalités et sont diffusées dans la salle par de nombreux haut-parleurs. Du point de vue du spectateur, chaque phrase du texte est prononcée par un personnage différent, placé ou se déplaçant différemment des autres. Pour la transformation de l'identité de la voix, le metteur en son Baptiste Poulain fait remarquer qu'elle est plus naturelle quand le comédien et la machine jouent ensemble la nouvelle identité. Il s'agissait donc pour les comédiens de construire et de respecter une conduite de timbre et d'intonation contenant des points de discontinuité entre chaque phrase et, parfois, entre les mots.

À l'inverse de la démultiplication de la voix, la machine permet de fusionner deux voix. Dans la troisième partie de *Nous les vagues*, on assiste à la préparation d'un attentat par un couple fusionnel de terroristes. Les deux comédiens disent le même texte simultanément et la fusion est opérée grâce à une synthèse croisée en temps réel. Le public perçoit alors une voix hybride, androgyne et non identifiable, forme allégorique de l'anonymat du terrorisme. Ce procédé a eu comme conséquence indirecte de forcer les comédiens à la parfaite synchronie, exercice du chœur auquel ils se sont prêtés avec enthousiasme.

SUIVI DE VOIX

Le suivi de voix consiste à comparer en permanence un flux de parole *live* (performance) à un flux de parole préenregistrée (référence). Plus largement, le suivi de voix hérite du suivi continu de données temporelles et permet donc de suivre tout type de geste vocal, doté ou non de message sémantique. Il permet donc de suivre non seulement la parole, mais aussi le chant ainsi que tout

10 *Une année sans été* est un texte de Catherine Anne qui a été mis en scène par Joël Pommerat en 2014 (création Odéon; mise en son de François Leymarie).

son émis par l'appareil vocal. En pratique, une répétition / lecture est enregistrée (une seule fois suffit). Des événements de toute nature (sonore, vidéo, régie plateau ou autre) sont « séquencés » le long de cet enregistrement via n'importe quel environnement logiciel (séquenceurs audio ou vidéo). Puis, cet enregistrement et ses événements relatifs sont stockés conjointement dans l'outil du suivi de voix. En situation de jeu, le flux de parole *live* est suivi, c'est-à-dire que la référence est perpétuellement synchronisée à la performance. Métaphoriquement, la performance devient la tête de lecture de la séquence organisée préalablement. Ainsi, les différents événements prévus et composés apparaissent et accompagnent la performance en temps réel.

Les applications du suivi de voix au théâtre sont nombreuses. Même si elles ne participent pas directement, de manière audible, à la dramaturgie sonore, elles permettent néanmoins d'en faciliter l'acheminement et de renforcer l'écoute de la machine. Le suivi de voix permet tout d'abord le déclenchement automatique d'événements discrets de type « Go » (lumière, vidéo, son, machinerie); la simulation de *word spotting* (reconnaître certains mots dans un flux de parole); la simulation de la reconnaissance de parole; la synchronisation automatique de surtitres; la régie sonore automatique comme dans le cas de *Nous les vagues*; le déclenchement et la variation d'effets sur la voix ainsi que la lecture augmentée.

Le suivi de voix permet aussi de suivre de manière continue et instantanée le jeu du comédien : synchronisation instantanée d'un flux (sonore, vidéo ou autre) à la performance scénique; conversion spécifique de l'identité de la voix (un locuteur parle avec la voix préenregistrée d'un autre); accompagnement synchrone au chant (comme un karaoké, mais à l'envers). Dans *Mars-Watchers*, l'expérimentation se termine par un canon de Henry Purcell, chanté par une seule comédienne, Constance Larrieu. Les autres voix du canon sont alors synchronisées à son tempo grâce au suivi instantané de voix.

Enfin, adjoint aux outils d'analyse vocale cités précédemment, le suivi de voix permet l'estimation en temps réel de variations prosodiques relatives à l'interprétation : comparaisons instantanées de l'intonation, de l'intensité, du degré d'articulation, de la qualité vocale, du débit de parole, du *timing* local (prononciation), du tempo global; la déduction du changement de stratégie verbale relative à l'accentuation et aux pauses et, conséquemment, la prise de décision scénique en temps réel selon l'interprétation du comédien. Par exemple, le choix de la musique ou de l'ambiance lumineuse peut se faire en fonction de l'expressivité du comédien.

Durant la première partie de *Nous les vagues*, les voix des deux comédiens sont transformées en de multiples identités vocales variablement spatialisées. Tous ces changements / mouvements ont été composés préalablement en studio dans une partition bien réglée. Du point de vue de

la régie, cette partie, d'une durée de six minutes, recense cent trente-six changements (Go), qui doivent être synchrones au phonème près – un véritable cauchemar pour le régisseur. De plus, à mesure que le processus se déploie, une accumulation d'autres voix envahit la salle par l'ajout d'une bande sonore. Si bien que, à la fin de cette partie, le public et le régisseur ne peuvent plus distinguer les voix transformées des comédiens en temps réel des voix transformées constituant la bande. C'est pour ces raisons que le suivi de voix s'est imposé comme unique solution viable pour la régie automatique de cette partie. Délivré d'une tâche fastidieuse, le régisseur a pu se concentrer sur le mixage sonore en salle.

Que ce soit pour réaliser un masque ou un déguisement sonore, une voix off particulièrement inhumaine ou encore un effet de chœur à partir d'un seul comédien, la machine apporte aujourd'hui de nouvelles solutions créatives à des questions ancestrales qui habitent le théâtre. En effet, la qualité et la rapidité du traitement automatique de la parole permettent aujourd'hui au metteur en son de proposer des solutions capables de remplir des fonctions relatives à la dramaturgie sonore. Cet article a présenté différentes solutions développées à l'IRCAM qui ont déjà participé à la dramaturgie sonore de spectacles.

Pour autant, les solutions exposées ne constituent pas en elles-mêmes un intérêt artistique. Le terme « solution » indique d'ailleurs qu'il y a au préalable une question posée, un besoin ou une problématique donnée qui, elle, appartient au projet artistique. Précisons que ces solutions peuvent, bien entendu, suggérer les questions. Cette dialectique entre question artistique et solution technique montre aussi que l'on ne peut pas surimposer une technologie à un projet déjà finalisé. Il ne s'agit pas ici d'utiliser la diffusion / amplification pour pallier un problème d'intelligibilité en salle ou encore un problème de santé d'un comédien. Il s'agit bien ici de créer avec des outils qui, s'ils viennent à manquer, rendent l'exécution du projet impossible. Il est donc nécessaire que ces techniques soient intégrées très tôt dans le processus de création et, idéalement, dès la genèse du projet.

De plus, puisqu'il s'agit de voix augmentées, ces solutions ne dépossèdent aucunement les comédiens de leur travail et leur proposent, au contraire, de nouvelles contraintes à explorer. D'ailleurs, la plupart des comédiens sollicités se délectent de ces nouvelles contraintes considérées comme de nouveaux exercices d'acteur. L'exemple de la transformation de l'identité de la voix est probant : on ne percevra la voix d'une femme que si un comédien homme, dont la voix est féminisée, joue lui-même un personnage féminin, ou encore, si la trame narrative suscite un tel travestissement. À l'instar du maquillage, la technologie permettra, dans ce cas, d'aider le comédien à se travestir en femme en produisant un léger différentiel dans la perception de sa voix. Cela demande donc au comédien et au créateur sonore un certain type d'écoute, voire une nouvelle proximité.

Cela renouvelle aussi le travail des créateurs sonores ou metteurs en son qui se retrouvent à l'interface entre le metteur en scène et les comédiens. Afin de proposer des réponses clairvoyantes vis-à-vis de situations dramaturgiques, les créateurs sonores doivent être constamment en dialogue avec le metteur en scène. Un nouveau langage est à créer pour chaque projet, le metteur en scène bénéficiant alors de nouvelles cordes à son arc pour la constitution de son œuvre. Enfin, pour boucler la boucle, les désirs formulés et les questions posées par ces créateurs sont autant de défis pour les chercheurs de l'IRCAM, toujours avides de voir leurs recherches et leurs développements mis à l'épreuve du plateau.

BELLER, Grégory (2011), « Arcane d'*Un mage en été* », *Théâtre/Public*, n° 201, (« Le son du théâtre III : Voix Words Words Words », Jeanne Bovet, Marie-Madeleine Mervant-Roux et Jean-Marc Larrue [dir.], p. 100-102.

BELLER, Grégory (2010), « Espresso: Transformation of Expressivity in Speech », communication donnée dans le cadre du colloque Speech Prosody, Chicago.

BELLER, Grégory (2009a), « Analyse et modèle génératif de l'expressivité : application à la parole et à l'interprétation musicale », thèse de doctorat, Paris, Université Paris 6 – IRCAM.

BELLER, Grégory, (2009b), « Transformation of Expressivity in Speech », dans Sylvie Hencil (dir.), *The Role of Prosody in the Expression of Emotions in English and in French*, Berne, Peter Lang, p. 259-284.

BELLER, Grégory (dir.) (2007), *Context Dependent Transformation of Expressivity in Speech Using a Bayesian Network*, actes du colloque ParaLing du 3 août, Paris, IRCAM.

BELLER, Grégory (2005), « La musicalité de la voix parlée », master, Paris, Université Paris 8.

BELLER, Grégory et Georges APERGHIS (2011), « Gestural Control of Real-Time Concatenative Synthesis in *Luna Park* », <http://articles.ircam.fr/textes/Beller11b/index.pdf>

BELLER, Grégory, Thomas HUEBER, Diemo SCHWARZ et Xavier RODET (2006), « Speech Rates in French Expressive Speech », communication donnée dans le cadre du colloque Speech Prosody, Dresde.

BELLER, Grégory, Nicolas OBIN et Xavier RODET (2008), « Articulation Degree as a Prosodic Dimension of Expressive Speech », communication donnée dans le cadre du colloque Speech Prosody, Campinas.

BELLER, Grégory et Xavier RODET (2007), « Content-Based Transformation of the Expressivity in Speech », communication donnée dans le cadre du Congrès international des sciences phonétiques, 6-10 août, Sarrebruck.

BELLER, Grégory, Diemo SCHWARZ, Thomas HUEBER et Xavier RODET (2005), « A Hybrid Concatenative Synthesis System on the Intersection of Music and Speech », communication donnée dans le cadre des Journées d'informatique musicale, 2-4 juin, Maison des sciences de l'homme, Paris.

BELLER, Grégory, Christophe VEAUX, Gilles DEGOTTEX, Nicolas OBIN, Pierre LANCHANTIN et Xavier RODET (2008), « IRCAM Corpus Tools : système de gestion de corpus de parole », *Traitement automatique des langues*, vol. 49, n° 3, p. 77-103.

SCHWARZ, Diemo, Grégory BELLER, Bruno VERBRUGGHE et Sam BRITTON (2006), « Real-Time Corpus-Based Concatenative Synthesis with CataRT », communication donnée dans le cadre de Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06), 18-20 septembre, Montréal.

VEAUX, Christophe, Grégory BELLER et Xavier RODET (2008a), « IrcamCorpusTools: An Extensible Platform for Speech Corpora Exploitation », communication donnée dans le cadre du colloque Language Resources and Evaluation Workshop on Emotions, 28-30 mai, Marrakech.

VEAUX, Christophe, Grégory BELLER et Xavier RODET (2008b), « IrcamCorpusExpressivity: Nonverbal Words and Restructurings », communication donnée dans le cadre du colloque Language Resources and Evaluation Workshop on Emotions, 28-30 mai, Marrakech.