

Arcane d'*Un mage en été*

Dans l'art magique et la prestidigitation, l'arcane est l'explication gardée secrète du fonctionnement ou de la série de passes grâce auxquelles est réalisé un tour. Si la règle déontologique “du secret” est de mise dans ces domaines, nous allons tout de même vous révéler l'arcane sonore d'*Un mage en été*. Cette pièce, créée au festival d'Avignon le 20 juillet 2010, est un monologue d'une durée d'une heure et demie, issue d'une adaptation de l'œuvre d'Olivier Cadiot portant le même nom¹. Mis en scène par Ludovic Lagarde, le mage, interprété par Laurent Poitreneaux, propose aux spectateurs un voyage dans l'espace-temps dont le convecteur principal est le son, réalisé par Greg Beller et David Bichindaritz avec les outils de l'IRCAM². L'acteur, immobile au sein d'une scénographie minimale, fait pénétrer le spectateur dans des tableaux qu'il crée par sa voix, traitée en temps réel et accompagnée d'une “bande-son”³ en constante évolution. Un dispositif d'enceintes entourant le public permet une mise en espace de ces éléments. Dans l'expérience fondatrice du *Colonel des Zouaves* en 1997, première collaboration du duo Cadiot/Lagarde avec l'IRCAM, la schizophrénie du personnage était en partie signifiée par l'intervention perturbatrice de la machine. Dans cette nouvelle collaboration, il s'agit à l'inverse de faire disparaître la machine, afin de conférer au personnage des pouvoirs surnaturels. Après avoir présenté le dispositif de diffusion, nous décrirons les différentes sources qui y sont diffusées, c'est-à-dire la bande-son, la voix du comédien et ses traitements, puis nous indiquerons quelques perspectives de recherche qui ont été initiées par ce projet.

Le dispositif de diffusion

Le dispositif de diffusion du son comporte neuf enceintes. Deux sont situées en fond de scène pour créer classiquement des effets de lointain. Une est placée au-dessus de l'acteur, de manière centrale, pour amplifier sa voix tout en préservant la directionnalité de celle-ci. Les six autres sont disposées autour du public et permettent une spatialisation fine de tous les éléments sonores. Si un tel dispositif immersif est aujourd'hui classique pour les concerts de musique contemporaine, il reste assez rare dans des théâtres qui ne proposent, la plupart du temps, que deux enceintes en façade. Associé au SPAT, outil de spatialisation de l'IRCAM, il permet le (dé-)placement de plusieurs sources sonores dans tout le théâtre (et virtuellement, au-delà), ainsi que la modification de la perception de son acoustique.

La bande-son

La bande-son du spectacle consiste en un mixage et une mise en espace de nombreux extraits musicaux, de bruitages et d'ambiances sonores. Grâce au logiciel LIVE d'Ableton⁴, elle est actualisée en temps réel par le régisseur son qui dispose d'une pédale lui permettant de gérer une centaine de déclenchements synchrones avec le texte, ainsi qu'un clavier contrôlant plusieurs paramètres d'effets comme le filtrage ou la réverbération. Omniprésente, la bande-son suit les propos du mage et ne les précède pratiquement jamais. D'un point de vue dramatique, ce lien causal donne au mage le pouvoir de créer et de manipuler des paysages sonores, afin d'y plonger le spectateur. Le son étant associé à une gestuelle adéquate, le mage donne ainsi l'illusion de faire

¹ Olivier Cadiot, *Un mage en été*, Paris, P.O.L., 2010.

² Institut de recherche et coordination acoustique (créé en 1970 à Paris par Pierre Boulez)

³ Sur l'usage de ce terme, voir l'article de Guillaume Trivulce, « La bande-son au théâtre », in *Théâtre/Public* n° 199, *Le son du théâtre. 2. Dire l'acoustique*, mars 2011, p. 32-35.

⁴ Logiciel de production musicale pour l'écriture, la composition, l'enregistrement.

apparaître ou d'attraper des sources dans des espaces sonores qu'il crée au fur et à mesure de son récit.

Le traitement de la voix

Amplification et travail sur la sensualité de la voix

La voix de l'acteur est constamment amplifiée, ce qui permet à celui-ci d'utiliser toute la dynamique de son appareil vocal allant du chuchotement au cri tout en préservant l'intelligibilité. De plus, un travail spécifique a été mené de manière à rendre la voix du mage tantôt rassurante et ronde, tantôt stridente et froide. Pour cela, nous avons utilisé un pré-ampli Avalon et de nombreux filtrages, auxquels viennent s'ajouter différents types de diffusion relatifs à la mise en espace de la voix.

Mise en espace de la voix

La notion d'espace étant complexe, il a fallu décrire cette notion de manière plus précise avec le metteur en scène. En effet, la mise en espace de la voix est relative à la position de l'acteur dans l'espace physique (la scène), à l'espace acoustique éventuellement modifié dans lequel on diffuse sa voix, enfin à la place que l'on donne à cette dernière dans cet espace. Grâce au SPAT de l'IRCAM, l'acoustique de la salle est virtuellement modifiable ainsi que la position des sources sonores dans cet espace virtuel. Dans *Un mage en été*, l'acteur restant physiquement au centre de la scène durant toute la pièce, plusieurs types de rapport entre l'espace acoustique recréé et la position de sa voix ont été explorés. Ainsi, la réverbération de la salle est constamment et dynamiquement modifiée pour figurer les lieux dans lesquels le mage se déplace (en extérieur, dans une cage d'escalier, dans des thermes romains...). De même, sa voix quitte parfois son corps pour faire le tour des spectateurs ou bien pour apparaître plus proche ou plus lointaine de ceux-ci. Par exemple, lorsque le mage mentionne *La Splendeur des Amberson*,⁵ sa voix est diffusée dans tous les haut-parleurs, renforçant artificiellement sa présence et donnant une perception de proximité au spectateur. Enfin le déplacement relatif d'autres sources sonores (bande-son) dans le même espace que sa voix donne l'illusion que l'on (qu'il) se déplace dans des paysages sonores.

Transformations réalistes de la voix

En tant que source sonore, la voix peut être traitée par de multiples effets (*delay / flanger / chorus / phaser*, ring modulation, vocodeur, transpositeur, granulateur...) qui ne vont pas forcément tenir compte de ses spécificités et altéreront son aspect naturel. Les sonorités ainsi produites sortent du contexte de la production vocale, nous les appelons alors des transformations non réalistes. Par exemple, une voix peut être démultipliée afin de créer des effets de groupe de manière synchrone ou asynchrone (effets de chœur, effet de foule, accumulation, *sampling*, dédoublement...).

Par opposition, une transformation réaliste de la voix produit des sonorités que notre perception est capable d'attribuer à une source vocale. En effet, le son de la voix véhicule plusieurs niveaux d'information dont l'identité du locuteur (sexe, âge, état de santé, taille, provenance...), l'expressivité (expression d'un état interne inaccessible à autrui), l'aspect sémantique et pragmatique (s'il s'agit de parole), la modalité et le style de parole⁶. Modifier chacun de ces niveaux d'information tout en gardant l'aspect naturel de la parole transformée (ou synthétisée) est l'un des buts que s'est donné l'équipe de recherche sur la voix de l'IRCAM⁷. Pour cela, il est nécessaire de pouvoir agir de manière indépendante sur différentes caractéristiques acoustiques de

⁵ *The Magnificent Ambersons*, film d'Orson Welles (1942).

⁶ Voir Greg Beller10b, 09a et 09c

⁷ Voir Greg Beller10b, 09b, 09c, 08c, 07a, 07b, Veaux08a.

la voix, comme son timbre et sa prosodie (intonation, intensité, débit de parole, degré d'articulation et qualité vocale⁸). Après plusieurs années de recherches et de développement, nous sommes aujourd'hui capables d'analyser et de transformer certaines de ces caractéristiques "en temps réel", c'est-à-dire de manière synchrone avec la voix du locuteur, tout en assurant une qualité permettant une utilisation dans un cadre artistique.

Transformations de l'identité de la voix

Un des vœux de l'auteur et du metteur en scène était de donner au mage le pouvoir de prendre les identités vocales de personnes identifiées, comme Proust, Nietzsche ou Crowley⁹ ou de personnes non identifiées (sa voix devient celle d'une femme). En ce qui concerne les trois personnalités mentionnées, malheureusement pour nous, il n'existe pas d'enregistrements. En nous appuyant sur des enregistrements anciens tirés d'une base d'archives sonores du XXe siècle (celle du CDMC, Centre de documentation de la musique contemporaine), nous avons réalisé ces transformations de manière empirique, en utilisant la combinaison d'un changement de timbre qui nasalise la voix et d'un effet radiophonique qui la situe dans une époque antérieure. Pour la voix de femme, nous avons légèrement élevé le timbre (ce qui produit l'effet d'un rétrécissement du conduit vocal) et avons travaillé avec l'acteur pour créer, de concert, une voix androgyne.

Duplication de la voix

Lors d'un passage dans lequel le mage est envahi par la magie noire, une voix plaintive émane de lui et se détache progressivement. Cette deuxième voix est générée par transformation de celle de l'acteur, grâce à un moteur permettant d'allonger automatiquement les voyelles. Couplage d'une détection de voisement et d'un modificateur de débit de parole, ce moteur permet d'accrocher des voyelles et de les transposer, produisant un pont entre la parole et le chant. Dans notre cas, nous ne l'avons pas utilisé pour générer du chant (avec des notes fixées) mais pour générer une litanie plaintive dont les variations sont issues d'une étude sur de la parole triste.

Transformations non réalistes, imaginaires et figures sonores poétiques

Un autre pouvoir que le metteur en scène souhaitait conférer au mage était la capacité de (se) télétransporter dans l'espace-temps. Il s'agissait de créer des effets illustrant le transport, comme un réseau d'ondes que le mage pourrait générer par sa voix et emprunter. Trois types d'effets ont été créés pour cela ; chacun intervient deux fois dans la pièce. Le premier est constitué d'un *delay* dont les multiples réflexions sont transposées, produisant des échos qui évoluent constamment vers l'aigu ou vers le grave. Par exemple, lorsque le mage s'élève au-dessus du sol pour voir les choses d'en haut, à l'instar d'un travelling arrière de *google-map*, sa voix s'accompagne d'échos toujours montants. Une fois qu'il est là-haut, les graves de sa voix sont filtrés et les aigus passent dans une réverbération donnant l'impression qu'il parle depuis les nuages. Le deuxième effet créé pour simuler le transport est un nouveau type de réverbération appelé réverbération à persistance harmonique. Cet effet accroche les partiels proéminents du spectre et les prolongent artificiellement, produisant des notes émanant de la voix. Finalement, cet effet a été utilisé une fois pour illustrer le déplacement de particules, et une autre fois, pour une mise en abîme dans laquelle le mage parle d'un harmonium qui accompagne la voix. Le troisième effet réalisé pour illustrer le transport est constitué d'une réverbération/écho qui ne prolonge que les transitoires, c'est-à-dire les consonnes. Il est utilisé une première fois pour créer une figure poétique. Le mage prononce les mots suivants : « Comme un vol de colombes désarticulées ». La structure consonantique est forte

⁸ Voir Greg Beller 08a et 06a

⁹ Aleister Crowley est un écrivain et occultiste britannique (1875-1947)

avec de nombreuses plosives¹⁰. Cette structure est prolongée, l'espace d'un instant, par cet effet qui donne alors à entendre l'envol des colombes. A la fin du spectacle, ce même effet dont les aigus sont cette fois-ci filtrés permet de donner l'illusion que le mage plonge dans l'eau, tintant sa voix d'échos obscurs et sous-marins.

En marge du mage

Plusieurs thématiques de recherche sur la voix ont été investies lors de la création d'*Un mage en été*. Même si celles-ci n'ont pas été directement exploitées dans le spectacle final, nous les évoquons car elles ont participé au processus de création.

Synthèse de la voix

Lors de sa première visite à l'IRCAM, Olivier Cadiot a été fasciné par le processus de synthèse à partir du texte¹¹ (il s'est d'ailleurs emparé du vocabulaire des chercheurs, qui se retrouve par endroits dans *Un mage*). Il a alors exprimé le souhait de pouvoir utiliser cette technologie chez lui, pour l'accompagner dans son processus d'écriture. Son vœu était de pouvoir soumettre un texte à la machine afin que celle-ci le dise avec la voix de Laurent Poitreneaux, comme si l'acteur lui-même en faisait la lecture. Nous avons donc enregistré le comédien et modélisé sa voix afin de pouvoir synthétiser n'importe quel texte avec son identité vocale et son jeu d'acteur. Ce travail a contribué au processus d'écriture, mais aucune voix de synthèse n'est audible dans le spectacle.

Suivi de parole

Le créateur vidéo de la pièce, Cédric Scandella désirait pouvoir déclencher des éléments visuels de manière synchrone avec le texte. Nous avons proposé l'utilisation du "suivi de parole", projet de recherche que nous sommes actuellement en train de développer. Le suivi de parole est une nouvelle technique permettant de comparer en temps réel un flux de parole à un enregistrement préalable de celle-ci. Beaucoup plus robuste que la reconnaissance de parole, le suivi de parole permet de nombreuses applications comme la régie synchrone au texte, la conversion de voix ou encore l'estimation du *timing* d'un comédien. Il est applicable dans le cadre de n'importe quelle performance vocale non improvisée. En voici son fonctionnement : tout d'abord, il nécessite l'enregistrement d'une répétition que nous appellerons « la référence ». Sur cette référence et grâce à n'importe quel éditeur de son sont inscrits des marqueurs relatifs à des actions que l'on souhaite synchrones avec le texte (allant de l'ouverture des rideaux au déclenchement de sons en passant par la modulation de paramètres de transformation...). N'importe quelle action discrète ou continue peut être reliée à la référence dans un travail de préparation de la performance. En situation de jeu, le suivi de parole va comparer sans arrêt la performance et la référence et déclencher les différentes actions prévues de manière synchrone avec la performance. De plus, cette comparaison acoustique entre ce qui est en train de se produire et ce qui s'est déjà produit permet d'accéder aux variations de la performance vis-à-vis de la référence. Ainsi, on accède au *timing* intrinsèque du comédien qui, en situation de jeu, peut ralentir ou accélérer par rapport à la référence. Cet écart temporel estimé en temps réel et de manière continue peut alors servir à ralentir/accélérer la bande-son de façon à ce qu'à un instant de la bande-son corresponde toujours le même mot. De même, on peut évaluer la variation prosodique du comédien par rapport à la référence pour paramétrer des transformations ou encore contrôler un flux vidéo. Si cette technique, encore dans une phase de recherche, n'a pu être utilisée dans *Un mage en été*, elle porte en elle suffisamment d'espoir pour que nous la relations.

¹⁰ Les plosives ou occlusives sont les consonnes que l'on prononce en fermant la bouche puis en l'ouvrant brusquement, comme "P" et "B".

¹¹ Voir Greg Beller 1b

Grâce aux divers procédés sonores dévoilés dans cet article, *Un mage en été* parvient à faire disparaître la machine malgré l'omniprésence de celle-ci. Un tel tour de magie a été notamment possible grâce aux périodes de travail communes avec le comédien qui s'est approprié les différents traitements que nous lui proposons et les a intégrés dans son jeu. Les nouvelles technologies de l'IRCAM ayant été mises au service de la mise en scène, cette réalisation propose au théâtre un nouveau rapport au son. Comment introduire la cinématographie par le son et non par l'image, comment décorporaliser la voix d'un comédien ou changer sa présence vocale de manière imperceptible sont autant de questions auxquelles *Un mage en été* fournit des éléments de réponse.

Bibliographie:

- [Beller11b] Beller, G., Aperghis, G., « Gestural Control of Real-Time Concatenative Synthesis in Luna Park », P3S, International Workshop on Performative Speech and Singing Synthesis, Vancouver, 2011, pp. 23-28
- [Beller10b] Beller, G., « Espresso: Transformation of Expressivity in Speech », Speech Prosody, Chicago, 2010
- [Beller09a] Beller, G., « Analyse et Modèle Génératif de l'Expressivité : application à la parole et à l'interprétation musicale », Thèse de Doctorat, Paris 6 - IRCAM, 2009
- [Beller09b] Beller, G., Veaux, C., Degottex, G., Obin, N., Lanchantin, P., Rodet, X., « IRCAM Corpus Tools: Système de Gestion de Corpus de Parole », TAL, Juillet 2009
- [Beller09c] Beller, G., « Transformation of Expressivity in Speech », The Role of Prosody in the Expression of Emotions in English and in French, ed. Peter Lang. (Peter Lang), 2009
- [Veaux08a] Veaux, C., Beller, G., Rodet, X., « IrcamCorpusTools: an extensible platform for speech corpora exploitation », LREC, Marrakech, 2008
- [Beller08a] Beller, G., Obin, N., Rodet, X., « Articulation Degree as a Prosodic Dimension of Expressive Speech », Speech Prosody 2008, Campinas, 2008
- [Beller08c] Beller, G., Veaux, C., Rodet, X., « IrcamCorpusExpressivity: Nonverbal Words and Restructurings », LREC workshop on emotions, 2008
- [Beller07a] Beller, G., Rodet, X., « Content-based transformation of the expressivity in speech », ICPhS, Allemagne, 2007
- [Beller07b] Beller, G., « Context Dependent Transformation of Expressivity in Speech Using a Bayesian Network », ParaLing, 2007
- [Beller06a] Beller, G., Hueber, T., Schwarz, D., Rodet, X., « Speech Rates in French Expressive Speech », Speech Prosody 2006, Dresden, 2006
- [Beller05a] Beller, G., « La musicalité de la voix parlée », Paris 8, 2005. [maîtrise de musique]