Master professionnel

Mention : Information et communication

Spécialité : Communication Médias

Option : Médias et management

La place de la création radiophonique sur les enceintes connectées

L'altération de la voix en contexte de virtualité numérique

Responsable de la mention information et communication
Professeure Karine Berthelot-Guiet

Tuteur universitaire : Valérie Patrin-Leclère

Nom, prénom : MELLOUET Clément

Promotion : 2017-2018

Soutenu le : 21/09/2018

Mention du mémoire : Très bien

École des hautes études en sciences de l'information et de la communication – Sorbonne Université
77, rue de Villiers 92200 Neuilly-sur-Seine I tél. : +33 (0)1 46 43 76 10 I fax : +33 (0)1 47 45 66 04 I celsa.fr

Remerciements.

Ce mémoire de fin d'étude n'aurait pu voir le jour sans le concours de nombreuses personnes, que je tiens à remercier.

Valérie Patrin-Leclère et Marc H'Limi, mes rapporteurs pour ce travail, ont su me guider, me suivre et m'apporter des conseils à chaque étape de sa réalisation. L'ensemble de l'équipe pédagogique du CELSA Sorbonne Université, représentée par Karine Berthelot-Guiet en sa qualité de directrice, a durant mes années passées dans cette école forgé mon esprit et mes connaissances de futur professionnel.

J'exprime toute ma gratitude aux membres du groupe radio RTL qui ont accepté de se confier dans le cadre d'entretiens essentiels à cette recherche : Rafaële Dhuicque-Filloux, Guillaume Duhamel, Jérémy Valet, Nicolas Dot, Pierre Laurent, Julien Le Ny, Alban Rivoal, Vladimir Bobrow et Sacha Ephraim. Je remercie également ceux, plus discrets, dont le nom ne figure pas sur cette page mais qui n'ont cessé de me transmettre leur connaissance et leur passion pour le média radio au quotidien.

Sommaire.

Remerciements. **2**

Introduction. **6**

I — Voix et virtualité : l'aboutissement d'un processus d'immersion et d'interaction. **18**

A. LES PROPRIÉTÉS DE LA VOIX : ENTRE MUSIQUE DU CORPS ET ARBITRAIRE DU SIGNE LINGUISTIQUE. 19

1. La musicalité de la voix, empreinte des origines sensibles de l'Homme. 19

2. La voix, support et médiation d'un arbitraire du signe. 24

3. La vive voix, incarnation de la personnalité. 26

B. LA VOIX À L'ORIGINE DES EFFETS PROPRES AU MÉDIA RADIO : L'AUTHENTICITÉ D'UN DISPOSITIF ACOUSMATIQUE. 28

1. La voix, condition de possibilité des effets du média radio : l'authenticité en tant que reflet des sens. 28

2. La voix de radio, travail de l'authenticité comme une relation de proximité. 34

C. LA VOIX, UNE OPPORTUNITÉ DE CRÉER DE NOUVELLES SITUATIONS D'INTERACTION ET D'IMMERSION AVEC LES PUBLICS DANS UN CONTEXTE DE VIRTUALITÉ NUMÉRIQUE. 38

1. La voix en contexte de virtualité : l'aboutissement des mécanismes d'immersion et d'interaction. 38

2. Un contexte porteur : entre « hyperesthésie » et saturation visuelle. 42

II — L'économie de la voix numérique : champ concurrentiel élargi et altération artificielle. **47**

A. LE MARCHÉ DES ENCEINTES CONNECTÉES : L'INNOVATION MÉDIATIQUE SOUMISE À LA DOMINATION TECHNOLOGIQUE DES GAFAM. 48

1. État des lieux des acteurs en présence dans le marché des enceintes connectées.48

2. Analyse des usages des enceintes connectées : des nouvelles situations d'écoutes pour la radio. 57

B. ANALYSE DE L'OFFRE DES MÉDIAS FRANÇAIS PRÉSENTS SUR L'ENCEINTE CONNECTÉE AMAZON ECHO. 61

1. Mise en contexte : une stratégie de déploiement régie par la notion de « média global ». 61

2. Analyse de l'offre des médias présents sur Amazon Echo en France. 64

C. DE LA RADIO À L'ASSISTANT VOCAL : LA DÉSINCARNATION DE LA VOIX COMME ENJEU TECHNOLOGIQUE. 76

1. L'altération de la vive voix dans son transport technologique : un processus déjà effectif en radio. 77

2. La voix artificielle réduite à un déterminisme technologique : la quête d'une qualité naturelle dans les logiques de conception des enceintes connectées. 80

III — Les enceintes connectées, indice d'une redéfinition du média radio entre menaces et opportunités. 90

A. LA PLACE DES VOIX DE RADIO AU SEIN DE L'ÉCONOMIE NUMÉRIQUE : ENTRE AVANTAGE CONCURRENTIEL ET PERTE D'IDENTITÉ. 91

1. L'économie de la voix numérique : des propriétés d'immersion et d'interaction intégrées à un champ concurrentiel élargi par la possibilité d'une incarnation. 91

2. La remédiation du média radio dans l'économie numérique de la voix : la place des stations historiques au sein de l'audio numérique. 94

B. LA PRAGMATIQUE RADIOPHONIQUE À L'ÈRE NUMÉRIQUE : LES ENCEINTES CONNECTÉES, INDICE D'UNE REDÉFINITION DES MÉTIERS DE RADIO. 100

1. La prégnance de la voix comme élément de définition d'une « pragmatique radiophonique ». 102

2. La perception des enceintes connectées comme un objet de redéfinition des métiers de radio. 111

C. RADIO DÉCADENTE OU RADIO RENOUVELÉE ? PERSPECTIVES DE POSITIONNEMENT DES STATIONS HISTORIQUES SUR LES ENCEINTES CONNECTÉES. 124

1. Les stations de radio au sein d'un champ concurrentiel élargi : une tension entre l'identité vocale et l'image de marque. 124

2. L'artificialisation des pratiques médiatiques : incorporer un savoir-faire technologique dans la création radiophonique. 126

Conclusion. **131**

Bibliographie & Sitographie. **136**

Annexes. **146**

Résumé. **223**

Mots-clés. **224**

Introduction.

DÉFINITION DE L'OBJET D'ÉTUDE

Par ce présent mémoire, nous souhaitons étudier la redéfinition du média radio dans une économie où son travail historique de la voix et du son se trouve à la fois concurrencé et accaparé par d'autres acteurs, en modifiant le sens et la réception. L'étude de ces processus est resserrée au cas du marché des enceintes connectées, qui altèrent le matériau vocal de deux manières : en mettant dans une situation d'horizontalité des contenus audio produits par une pluralité d'acteurs pour lesquels la voix n'est pas liée de la même manière à leurs identités de médias, et en déplaçant l'expression vocale, phénomène humain et corporel que la radio transforme tout en en préservant l'authenticité, dans une Intelligence Artificielle qui l'incarne avec une personnalité construite de toutes pièces.

En d'autres termes, notre objectif est d'observer la remédiation[1] des formes radiophoniques dans un environnement numérique — leur transformation et leur réactualisation — et de définir la place des stations de radio françaises — leurs avantages concurrentiels, leurs risques, et la manière dont ceux-ci sont appréhendés au sein même des stations — sur les enceintes connectées. Ce travail devrait par ailleurs nous permettre d'évaluer la prégnance de deux notions dans ce processus : celle de média global, soit la tendance à l'ubiquité, à la multiplication des supports de prise de parole caractéristique des marques médias sur le numérique, et celle d'artificiel, illustrée par la voix de l'enceinte connectée, reliée à un algorithme, qui semble être la voie majeure aujourd'hui empruntée pour paradoxalement personnaliser l'offre médiatique à chaque utilisateur.

Notre recherche comporte un fort caractère exploratoire. En effet, le marché des enceintes connectées demeure embryonnaire à l'échelle internationale, avec des degrés de pénétration distincts selon les territoires. Si les États-Unis, berceau de la technologie, présentent une forte adoption des publics, la France apparaît en retrait. En traitant tour à tour de plusieurs aspects de l'environnement médiatique que les enceintes connectées construisent, nous souhaitons ainsi dépeindre un contexte dans lequel les stations de radio s'insèrent, entre menaces et opportunités pour leur média. Dans ce contexte, cette photographie devrait nous permettre d'établir des recommandations pour les stations de radio en termes de positionnement.

---

[1] Bolter, J.D., Grusin, R., *Remediation. Understanding New Media*, MIT Press, 1998.

INTÉRÊT UNIVERSITAIRE & PROFESSIONNEL DU SUJET

La radio est un média longtemps « délaissé »[2] par la recherche que des groupes scientifiques tels que le Groupe de Recherches et d'Etudes sur la Radio (GRER) tentent de revaloriser[3] — du moins dans les territoires de la francophonie — dans un contexte où elle se découvre une nouvelle actualité avec les modes de consommation numériques. Nous pensons ainsi que cette recherche, qui traite de la radio sous l'angle d'une actualité que nous considérons la concerner au premier plan car usant de la voix, s'ancre dans un domaine renouvelé de la recherche en Sciences de l'Information et de la Communication. Nous questionnons l'identité d'un média, et la signification des changements qu'il traverse pour l'avenir des pratiques et des modes de consommation médiatiques. La question centrale de notre étude est en effet celle du corps, incarné par une voix qui au gré des évolutions techniques se déplace dans des ondes radios, sur des pages web, jusqu'à être remplacée en sa source même par une Intelligence Artificielle.

La mise en rapport des technologies vocales avec le statut et le positionnement du média radio présente également un intérêt professionnel, et ce à deux degrés : observer la redéfinition de la pratique des métiers de radio, déjà à l'œuvre avec les réseaux sociaux par exemple, et tenter de définir le positionnement qu'elle pourrait adopter face à l'ouverture du nouveau marché médiatique que constituent les enceintes connectées.

CONTEXTE & CADRE DE L'ÉTUDE

Interviewé par *Le Figaro* en avril 2017, Mathieu Gallet, alors Président de Radio France, déclarait que « la prochaine grande révolution dans le monde des médias sera la voix »[4]. La conquête de la voix sur le numérique a en réalité déjà commencé, et ce en deux mouvements : le renouveau de l'engouement pour les podcasts et la commercialisation des enceintes connectées, qui ont tous deux émergés à partir de l'année 2014 dans le marché nord-américain.

---

[2] Tétu, J.F., « La radio, un média délaissé », *Hermès, La Revue*, 2004/1 (n° 38), p. 63-69. Disponible à l'adresse suivante : https://www.cairn.info/revue-hermes-la-revue-2004-1-page-63.htm

[3] Antoine, F., et. alii, *Analyser la Radio. Méthodes et Mises en Pratique*, éd. De Boeck, 2016.

[4] *Le Figaro*, « Mathieu Gallet : « La prochaine révolution numérique sera la voix », 25/04/2017. Disponible à l'adresse suivante : http://www.lefigaro.fr/medias/2017/04/23/20004-20170423ARTFIG00111-mathieu-galletla-prochaine-revolution-numerique-sera-la-voix.php

Les podcasts sont en train de construire leur légende. *Serial*[5], lancé en 2014 aux États-Unis, apparaît comme la pierre angulaire de la « renaissance du podcast »[6]. Les podcasts, en facilitant à la fois l'accès et la création de contenus audio, destinés à être consommés dans une situation de mobilité (comme l'induit le préfixe *-pod*, provenant de l'Ipod, baladeur intelligent conçu par Apple), sont un élément décisif dans l'ouverture de l'économie de la voix à une variété d'acteurs sur le numérique. En ce qu'ils sont des contenus sonores, ils attestent de la première perte de monopole du média radio dans le travail de la voix, et ce deux manières : lorsque d'autres acteurs médiatiques s'en emparent dans une logique de média global, caractéristique de l'économie numérique, visant à multiplier les supports de prise de parole ; et lorsque des amateurs créent leurs propres contenus, parfois avec plus de succès que des acteurs professionnels. Ce phénomène est qualifié par Henry Jenkins de *convergence* des pratiques médiatiques[7]. L'offre de podcasts se répartit en deux grandes catégories. Les podcasts natifs — uniquement dédiés à une consommation digitale — sont produits par une diversité d'acteurs, allant des radios (notamment France Culture) aux amateurs et aux entreprises de presse, qui peuvent y trouver une importante source de revenus. Pour exemple, 25% des revenus de *Slate*, *pure player* traitant de l'actualité politique et culturelle, proviennent de ses vingt quatre podcasts[8]. Les podcasts de rattrapage constituent l'autre partie de l'offre, prisée par les radios, qui y proposent leurs contenus en *replay*, à la manière de la télévision. Ces contenus délinéarisés (car non proposés de manière linéaire, en direct) sont aujourd'hui un succès international, pénétrant peu à peu le grand public. En 2018, 73 millions d'Américains écoutent au moins un podcast par mois[9].

Les enceintes connectées sont arrivées sur le marché nord-américain en 2014 avec Amazon, alors précurseur. Ses enceintes Echo, dont plusieurs déclinaisons sont désormais disponibles, sont animées par une Intelligence Artificielle nommée Alexa. Celle de Google, sobrement appelée Google Assistant et matérialisée dans l'enceinte Google Home, n'est apparue qu'en 2016, et le HomePod d'Apple n'a été lancé qu'en février 2018 aux États-Unis. A noter qu'Apple a été le premier de ces GAFA (acronyme désignant les géants américains du numérique, Google, Amazon,

---

[5] Créé en 2014 aux États-Unis, *Serial* relate un crime véridique sous l'angle du journalisme d'investigation. Ce podcast est produit par *This American Life*, studio auteur d'un programme éponyme et de *S-Town*, eux aussi de grands succès.

[6] *New York Magazine - Daily Intelligencer*, « What's Behind the Great Podcast Renaissance? », 31/10/2014. Disponible à l'adresse suivante : http://nymag.com/daily/intelligencer/2014/10/whats-behind-the-great-podcast-renaissance.html

[7] Jenkins, H., *Convergence Culture: Where Old and New Media Collide*, New York University Press, 2006.

[8] *Digiday*, Moses, L., « 'We're Pivoting to Words: *Slate* Says Podcasts Are Now 25 Percent of Revenue », 16/01/2018. Disponible à l'adresse suivante : https://digiday.com/media/slate-pivot-to-words/

[9] Edison Research, « The Podcast Consumer 2018 », 19/04/2018. Disponible à l'adresse suivante : http://www.edisonresearch.com/podcast-consumer-2018/

Facebook et Apple, auquel on ajoute parfois Microsoft, devenant alors GAFAM) à créer un assistant vocal pour le grand public. SIRI, son logiciel de commande vocale, était disponible dès 2011 dans ses iPhone 4S. En France, les sorties de ces différents modèles d'enceintes connectées ont été plus concentrées dans le temps : le 3 août 2017 pour Google, les 13 et 18 juin 2018 pour Amazon et Apple. Cet objet propose une interaction directe entre un utilisateur et une Intelligence Artificielle grâce à la voix. Cette dernière ne répond plus uniquement à la consommation d'un contenu spécifique, comme avec le podcast, mais à une immersion du numérique, de la virtualité, dans la vie de son utilisateur. L'enceinte connectée propose des fonctions plurielles — la connexion à l'ensemble domotique d'un foyer, l'achat en ligne, l'écoute d'un contenu audio, etc. — uniquement gérées par l'ouïe et la parole. Ces modes de compréhension et d'expression, en ce qu'ils émanent directement du corps, effacent la distance qui subsiste en une personne et son écran, que la Réalité Virtuelle essaie de conjurer avec un casque, au plus près des yeux. Avec la voix, la relation est immédiate, tant du point de vue de l'espace que de celui du temps.

Sorte de transistor augmenté, tant par son aspect que par la relation qu'elle propose à l'usager qu'elle complète avec la possibilité d'une interaction, l'enceinte connectée est un objet rassemblant des fonctions pratiques (la météo, l'information routière, la possibilité de faire sa liste de courses ou d'allumer les lumières de son salon sans se lever de son canapé) et l'accès à une diversité de contenus audio dans une situation d'écoute auparavant réservée à la radio historique. Ces derniers rassemblent de la musique, des podcasts, et des contenus créés spécifiquement pour ces enceintes soit en utilisant sa technologie de *text-to-speech* (les mots écrits d'un article sont lus à haute voix par l'Intelligence Artificielle), soit en formatant des contenus courts, adaptés à ce support. La radio, média historique du travail de la voix, qu'elle met en ondes, occupe une place singulière dans ce contexte.

PROBLÉMATIQUE ET HYPOTHÈSES DE TRAVAIL

Problématisation du sujet

Avec les enceintes connectées, la radio se retrouve dans une position similaire et concomitante à celle qu'elle occupe dans le marché des podcasts : favorisée par l'usage de la voix, qui est sa matière première, mais concurrencée par une multiplicité d'acteurs obéissant tous à la

logique d'un média global, déployant un maximum de canaux de distribution de ses contenus et de sa marque. Nous pourrions identifier dans ce processus une perte de spécificité, un effacement des particularités des médias au profit d'un idéal d'ubiquité à la fois recherché et contraint par l'économie numérique, où la concurrence se densifie et où chaque support perd en rentabilité économique. Les enceintes connectées ont cependant de spécifique qu'elles semblent constituer l'ultime aboutissement de ce processus. Elles vont jusqu'à remplacer la source la voix, en la déplaçant d'un corps à une machine. Elles sont par conséquent le théâtre d'une redéfinition du statut de la voix en contexte médiatique et, par extension, de la raison d'être du média radio. En effet, la voix n'est plus seulement utilisée par des entreprises diverses pour produire du contenu, ce qui est la caractéristique saillante du marché des podcasts ; avec les enceintes connectées, sa source, son origine corporelle est altérée en se déplaçant d'un humain à une machine — d'un animateur de radio à un robot parlant. Les enceintes connectées constituent en cela un carrefour dont la radio est au centre, car ancrée dans le travail de la voix. Deux routes se croisent : l'aboutissement d'un processus d'immersion du numérique dans la réalité physique avec la voix, émanation corporelle support d'une interaction immédiate, et la progressive artificialisation de la création médiatique dans un marché numérique qui se caractérise par son déterminisme technologique. Sur ce point, David Mattin, Directeur du Département *« Trends & Insight »* du cabinet TrendWatching expliquait lors du salon The Next Web qui se tenait à Amsterdam en mai 2018 que les tendances ne se définissent plus avec l'observation des comportements des utilisateurs mais en évaluant la capacité d'une évolution technologique à répondre à un besoin. Les « assistants vocaux vont réussir à s'inscrire dans notre quotidien car ils répondent à « des besoins que l'on croit être spécifiques à la nature humaine ». Par exemple : une IA [Intelligence Artificielle] peut vous apporter du réconfort ou de la compagnie »[10]. Derrière l'usage d'une voix artificielle dans les enceintes connectées se tient l'entier problème que génère ce déplacement de l'homme à la machine. La voix, en tant qu'émanation du corps, est incarnée par celui qui la prononce ; incarner une voix à l'aide d'une machine revient à créer une personnalité artificielle — plus qu'un assistant, un compagnon —, qui remet en cause les fondements du dispositif radiophonique et des métiers qui le composent.

[10] Propos de David Mattin lors de sa conférence au salon The Next Web rapportés par *Numerama*, Turcan, M., « Comment les assistants vocaux vont devenir nos « compagnons de vie » », 26/05/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/378318-comment-les-assistants-vocaux-vont-devenir-nos-compagnons-de-vie.html

Problématique et hypothèses

La voix constitue par conséquent une grille de lecture pertinente pour observer la redéfinition de la radio sur le numérique. Concurrencée par une multiplicité d'acteurs dans la production de contenus vocaux, son matériau principal et objet de son savoir-faire se trouve désormais encadré — lorsque que des contenus radiophoniques sont relayés sur les enceintes connectées — voire remplacé — dans le cas du *text-to-speech* — par une voix artificielle. La problématique qui structurera notre travail est la suivante :

> **L'engouement des entreprises médiatiques pour les outils mobilisant la voix, dont les enceintes connectées constituent un aboutissement en contexte numérique, construit-il des offres au sein desquelles les acteurs historiques de la création radiophonique se fondent dans le champ plus large de l'audio numérique ?**

Notre réponse à ce questionnement sera structurée par trois hypothèses (1, 2 et 3) élucidées à plusieurs étapes de notre travail :

1. Les technologies vocales font l'objet d'un engouement des entreprises de médias émanant d'une double opportunité reflétée dans les discours d'escorte des acteurs qui s'y engouffrent : la possibilité d'atteindre les publics dans de nouvelles situations d'écoute, et la diversification d'une offre éditoriale avec un *médium* aux atours plus authentiques dans une économie de l'attention caractérisée par un effet de saturation visuelle.

2. Sur les enceintes connectées, l'authenticité permise par la voix est altérée par deux éléments concomitants : l'interaction qui s'instaure entre l'utilisateur et une Intelligence Artificielle d'une part, et la création de contenus médiatiques adaptés à cette relation d'autre part.

3. Entre remédiation et convergence, la radio est dans une position ambivalente dans le champ concurrentiel des enceintes connectées. Sa puissance historique et économique est contrebalancée par la menace d'acteurs plus petits mais dédiés à la création digitale et par le déplacement de la voix dans une personnalité artificielle, allant jusqu'à redéfinir la pratique des métiers de radio.

MÉTHODOLOGIE

Présentation du corpus et de son mode de traitement

Pour mener à bien notre recherche, nous avons décidé de conjuguer trois méthodes de travail ayant chacune pour objectif de résoudre un problème précis (1, 2 et 3). L'usage de ces méthodes sera marqué par le caractère exploratoire de notre recherche, et adopteront une démarche à la fois thématique et démonstrative. En effet, les enceintes connectées constituent en France un marché embryonnaire, dont notre souhait est de capturer une photographie en traitant des multiples aspects qui la composent. Une importante disparité entre les possibilités apportées par la conjugaison des propriétés de la voix et de la virtualité numérique sera perceptible au regard de l'offre médiatique existante en 2018. Notre objectif est, en répondant à notre problématique, d'observer le statut de la radio au sein d'une économie numérique où son travail de la voix est déplacé vers une multitude d'acteurs, menaçant son hégémonie, avant de formuler des recommandations à destination des stations de la bande FM qui se remédient dans ce contexte.

1. Une recherche théorique visant à définir les propriétés du matériau vocal, qui selon nous est à l'origine de l'authenticité ressentie à l'écoute du média radio, d'autant plus intense lorsque l'on respecte son dispositif d'origine, acousmatique (dénué de l'image de la source du son). Les travaux de chercheurs d'une pluralité de disciplines, allant des Sciences de l'Information et de la Communication dans le cas de *L'Analyse du Discours Radiophonique* de Patrick Charaudeau aux ouvrages de Marie-France Castarède (qui emprunte plusieurs de ses théories à la psychanalyse) ou d'Iván Fónagy, spécialisé en psycho-phonétique, ont été mobilisés. La constitution de ce corpus théorique a été facilitée par la lecture du manuel *Analyser la Radio. Méthodes et Mises en Pratique*, publié en 2016 aux éditions De Boeck[11], qui rassemble de manière synthétique les travaux d'importance pour une grande variété d'axes de recherche sur ce média. Il nous a semblé essentiel de mêler les apports de plusieurs disciplines car la voix a de spécifique son origine corporelle, qu'il nous faut d'abord comprendre pour pouvoir en étudier l'altération lorsqu'elle se déplace d'un sujet à un objet (l'enceinte connectée et son Intelligence Artificielle). Nous pensons ainsi qu'après avoir identifié les caractéristiques du matériau vocal, nous serons en mesure de définir la spécificité du média radio, puis d'en évaluer la redéfinition sur le numérique.

[11] Antoine, F., et. alii, *Analyser la Radio. Méthodes et Mises en Pratique*, éd. De Boeck, 2016.

2. L'analyse des discours d'escorte et de l'offre médiatique proposée sur Amazon Alexa en France, afin d'avoir une idée précise des positionnements des différents médias impliqués sur la plateforme, que nous avons détaillés plus finement encore pour l'offre des radios. L'arrivée d'Amazon Alexa, l'Intelligence Artificielle développée par la firme de Jeff Bezos, dans la gamme d'enceintes connectées Echo en France le 13 juin 2018 était un événement appuyé par la présence de nombreux médias y proposant des contenus gratuitement. Ces derniers, afin de profiter de cet effet d'annonce, ont publié des communiqués de presse ou des articles s'y apparentant car présentant leurs offres — appelées *skills*, équivalent chez Amazon des applications (ou *apps*), dénomination à laquelle nous sommes davantage habitués grâce au téléphone mobile. Notre analyse de l'offre médiatique proposée au lancement d'Amazon Alexa en France comprend deux éléments (*a* et *b*).

a) Les présentations des offres des médias associés au lancement d'Amazon Echo en France.

Notre corpus d'analyse se réduit à une liste de médias s'associant au lancement de l'enceinte connectée Amazon Echo en y proposant une *skill* dès sa sortie sur le marché français. Nous avons analysé pour chacun d'entre eux un contenu de présentation de leur offre publié entre le 1er et le 30 juin 2018 — encadrant ainsi le lancement d'Amazon Alexa d'une quinzaine de jours en amont et en aval — sur le web. Par ces contenus de présentation de leurs offres, nous entendons plus précisément des communiqués de presse et/ou des articles sur les sites Internet des médias concernés ayant pour but de les exposer. Ces médias ont été identifiés grâce à des articles de la presse spécialisée en marketing et en communication les répertoriant (un article de *L'ADN*[12] et un de *CB News*[13]), puis les contenus de présentation de leurs offres recherchés sur leurs sites respectifs avec nos bornes temporelles en tête.

L'objectif de ce regroupement de médias distincts est comparatif, et son analyse qualitative : selon leur support d'origine (la presse, la télévision, la radio ou le numérique), mettent-ils en avant les mêmes arguments ? Quels sont les points communs ? Quelles sont les spécificités ? Au-delà de l'effet d'annonce dont chacun profite, et que les bornes temporelles choisies pour la constitution de notre corpus — quinze jours avant et après le lancement d'Amazon Echo — accentuent, nous

[12] *L'ADN*, Huot, A., « Lancement d'Alexa en France : quels médias ont répondu présent ? », 13/06/2018. Disponible à l'adresse suivante : http://www.ladn.eu/media-mutants/les-medias-francais-sur-alexa/

[13] *CB News*, Wojciak, T., « Amazon lance Alexa en France : les médias s'y engouffrent… », 06/06/2018. Disponible à l'adresse suivante : http://www.cbnews.fr/mobile/amazon-lance-alexa-en-france-les-medias-s-y-engouffrent-a1043088

voulons savoir si l'on observe une tendance à l'effacement des spécificités liées au support d'origine pour laisser émerger une marque média, ou si la particularité du support d'origine y est figurée, explicitement ou implicitement.

La finalité de cette étude comparative est d'évaluer la prégnance de la notion de média global dans les stratégies médiatiques à l'aune du lancement d'Amazon Echo, soit deux éléments : la démultiplication des supports de prise de parole comme perte de spécificité des atours de marques historiques ; le déploiement d'une stratégie multicanal contrainte par un support une nouvelle fois initié et dominé par les GAFAM (ici Amazon, mais Google et Apple, les deux autres marques occidentales ayant leurs enceintes connectées, en font également partie).

b) L'offre des médias radio français sur Amazon Alexa.

Nous avons par ailleurs répertorié les *skills* proposées par des stations de radio sur Amazon Alexa en France, à partir des résultats proposés à la recherche de l'occurrence « radio » sur amazon.fr/skills. Leurs offres ont été comparées de manière quantitative, en déterminant si certains types de contenus étaient proposés ou non dans leurs *skills* : le flux en direct, des contenus délinéarisés, et des contenus dédiés à Amazon Alexa. Notre objectif final est de déterminer si les radios françaises font l'effort de concevoir des contenus destinés à l'expérience dans laquelle est immergé un utilisateur d'enceinte connectée, ou si elles vivent ce support comme une opportunité d'accroître la portée de leurs contenus habituels.

Selon nos résultats, nous devrions être en mesure de constater si les enceintes connectées constituent pour les radios un enjeu d'expression et de figuration de leur identité vocale. Il nous faudra cependant nuancer ce constat avec la taille embryonnaire du marché des enceintes connectées en France et le coût que représente l'allocation de ressources à la création d'un contenu qui leur est dédié.

3. Une série de neuf entretiens semi-directifs avec des personnes travaillant au sein du groupe radio RTL, à l'antenne ou à la régie publicitaire. L'objectif : vérifier si, concrètement, la voix est pour eux un enjeu, à la fois dans leur travail quotidien et à la lumière de la promesse d'innovation qu'apportent les enceintes connectées. Ces personnes ont été interrogées sur une pluralité de

dimensions, dont l'importance diffère pour chacune selon son poste, comme en attestent les guides d'entretiens[14]. Deux points majeurs seront ainsi évalués :

- La prégnance de la voix comme élément de définition d'une « pragmatique radiophonique »[15], c'est-à-dire une vision du monde articulée autour du son, et ayant pour mécanismes principaux les notions de forme et de mise en forme présentes dans le processus de mise en ondes.
- La perception des enceintes connectées comme un objet de redéfinition des métiers de radio, car directement axé sur la voix dans un environnement numérique où ce matériau est accaparé par une multitude d'acteurs et est déplacé vers des technologies, symptôme d'une progressive artificialisation des pratiques médiatiques.

Le choix d'une focalisation sur des employés du groupe radio RTL répond à plusieurs éléments. Outre la diversité de ses actifs radio — la première radio de France, généraliste, ainsi que RTL2 et Fun Radio, deux stations musicales aux publics distincts —, le groupe RTL a fait la Une de l'actualité médiatique en parachevant sa fusion par absorption par le groupe M6 au sein d'un actionnariat global du groupe allemand Bertelsmann. Forts de cette actualité, nous supposons que les employés du groupe radio RTL ont à l'esprit une notion de changement, de transformation de leur métier *a priori* plus vive car prise au beau milieu d'une fusion d'ampleur, augmentant sa puissance sur le numérique et ajoutant des actifs TV à ses leviers d'action. Le fait que nous nous focalisions sur des employés de ce groupe est également lié à des contraintes matérielles, à un souci de facilité d'accès à une diversité de profils présentant des points de vue différents et/ou complémentaires dans un même contexte.

Les profils interrogés, s'ils diffèrent, présentent des similitudes nécessaires à la vérification des deux points d'analyse que nous avons évoqués plus tôt : le travail de la radio en tant qu'agencement d'un matériau vocal et sonore, direct dans le cas de producteurs (Julien Le Ny et Alban Rivoal, chargés de production chez Fun Radio), de régisseurs (Vladimir Bobrow et Sacha Ephraim, travaillant pour RTL), et de réalisateurs de messages radio (Rafaëlle Dhuicque-Filloux, co-fondatrice et directrice de ZéroDécibel, structure de production de spots publicitaires), ou indirect dans le cas des chargés de marketing radio et des chefs de projet d'Opérations Publicitaires Spéciales (OPS). Ces derniers ont été interrogés pour leur vue stratégique du secteur radio

[14] Les guides d'entretiens sont proposés avant chaque retranscription en Annexes 1, 4, 7, 9, 11, 13 et 15, pp.146-206.

[15] Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001, p.15.

(Guillaume Duhamel pour son expertise et sa connaissance du média, Jérémy Valet pour son point de vue stratégique sur la voix et ses évolutions, dont la création de M6 Unlimited, structure proposant des solutions *crossmedia* à des annonceurs qu'il co-dirige, constitue un indice). Les chargés de marketing ont été interrogés pour leur connaissance du média radio appuyée par des données à la fois quantitatives et qualitatives sur ses publics grâce à leur travail quotidien des données d'audiences. Par souci de former un tout cohérent, la mobilisation de chargés de marketing radio a été circonscrite à deux personnes présentant chacune un intérêt pour la voix ou les enceintes connectées, intégré directement à leurs fonctions. Pierre Laurent, en charge du Laboratoire Sonore, est spécialisé dans l'appréciation et la perception du son, qu'il utilise à des fins d'efficacité publicitaire ; Nicolas Dot est quant à lui chargé d'une veille sur les enceintes connectées, assurant de son savoir dans ce domaine et du fait que ce sujet soit une préoccupation pour son service.

Cette démarche, répartie en trois méthodes, sera complétée par l'utilisation d'articles de presse et d'études du marché des enceintes connectées, afin de nourrir notre réflexion d'apports concrets. C'est sur ce seul point que notre limitation de l'analyse au marché français pourra être contredite, car nombre de sources proviennent de pays anglo-saxons — en premier lieu des États-Unis, où les enceintes connectées sont nées et ont davantage intégré les usages des publics que dans l'hexagone.

ANNONCE DU PLAN

Afin de traiter notre problématique et ses hypothèses, nous effectuerons notre recherche en trois temps : l'analyse du matériau vocal (ses propriétés, son transport dans le média radio, et sa logique d'intégration dans la virtualité, domaine dont il partage les qualités d'immersion et d'interaction) ; l'étude de l'économie de la voix numérique (définie par l'intégration de ce matériau dans les stratégies de contenus d'une pluralité d'acteurs médiatiques, et par son appropriation par les concepteurs d'une voix artificielle) ; la redéfinition du média radio à l'aune de ces différents constats, évaluée de manière externe avec la notion de marque-média, puis de manière interne en confrontant les personnes constituant son savoir-faire aux évolutions de leur secteur d'activité, dont les enceintes connectées sont à la fois un indice et un aboutissement.

Notre recherche entreprendra par conséquent d'analyser un environnement encore en construction. A la fois contrainte et opportunité pour la radio, cette donnée temporelle sera particulièrement visible lors de l'analyse des contenus proposés en 2018 sur les enceintes connectées, peu développés en France et en phase d'amélioration aux États-Unis. Le caractère exploratoire de notre étude a selon nous de pertinent qu'il devrait permettre de définir le contexte numérique dans lequel les stations de radio se remédient, et apporter des pistes stratégiques en termes de positionnement. Il justifie ainsi que l'articulation de nos parties réponde à la conjugaison de trois méthodes, aspects, et thématiques, voués à composer une fresque d'un paysage médiatique que ses parties prenantes continuent de peindre : les propriétés de la voix pour les personnes qui la prononcent et qui l'écoutent dans une pluralité d'espaces (la conversation immédiate, la radio et la virtualité numérique), les logiques économiques et technologiques des acteurs qui vendent des enceintes connectées ou y proposent du contenu, et enfin la place du média radio tant du point de vue de son support que des marques et des savoir-faire qui le composent.

I — Voix et virtualité : l'aboutissement d'un processus d'immersion et d'interaction.

A la fois grille de lecture et objet concentrant l'ensemble de notre analyse, la voix est avec les enceintes connectées ce qui rassemble une pluralité d'acteurs : les concepteurs de technologies, les plateformes faisant office d'intermédiaires, les entreprises médiatiques, les utilisateurs. Depuis la sortie d'Amazon Alexa sur le marché américain en 2014, des études s'enchaînent pour promettre à la voix un futur doré. La raison ? L'interaction directe qu'elle permet entre un contenu et son consommateur. Qu'il s'agisse de commander en ligne, d'allumer les lumières de son salon, d'écouter de la musique ou la radio, l'utilisation est facilitée car elle ne mobilise que la voix. Mais la voix est également vectrice d'émotions très fortes, et source d'une attention de grande qualité de la part de ceux qui font le choix de l'écouter. Le succès des podcasts en est la preuve la plus édifiante. De ce constat provient une intuition, fil rouge de notre première partie : la voix est l'aboutissement de l'immersion des utilisateurs dans un Internet dit symbiotique — où la frontière entre réalité physique et monde virtuel continue de s'amincir — parce qu'elle est une projection du corps en un sujet qui en parlant s'inscrit dans le monde. Le matériau vocal a ainsi de paradoxal que son ancrage dans le corps parachève la virtualité avec les enceintes connectées.

Notre première partie commencera par tenter de définir les qualités intrinsèques de la voix, pour les lier aux fondements de l'effet d'authenticité du dispositif radiophonique, que le numérique vient à la fois altérer (lorsqu'il y ajoute de l'image) et prolonger (quand la voix devient également le mode d'accès au contenu, en amplifiant les effets d'immersion et d'interaction).

A. LES PROPRIÉTÉS DE LA VOIX : ENTRE MUSIQUE DU CORPS ET ARBITRAIRE DU SIGNE LINGUISTIQUE.

1. La musicalité de la voix, empreinte des origines sensibles de l'Homme.

Le terme *voix* désigne chez l'Homme un « ensemble de sons produits par la bouche et résultant de la vibration de la glotte sous la pression de l'air expiré ; faculté d'émettre ces sons »[16]. Elle est une émanation du corps — dont la bouche constitue l'ouverture — qui traduit un tempérament d'origine émotionnelle. La voix est utilisée pour communiquer un ensemble d'états aux autres et à soi-même, car son empreinte sonore inscrit la vie intérieure dans le monde. Marie-France Castarède, auteur d'un ouvrage à forte empreinte psychanalytique recouvrant de manière passionnante les caractères de la voix, qu'elle qualifie de *sortilèges* dans une tentative d'en retranscrire le mystère ontologique, traduit ce constat en ces termes :

> « Pour l'enfant de l'homme, la mise au monde est mise en voix : c'est le cri primal ou parole de vie... Sa voix ne s'éteindra que lorsqu'il rendra son ultime souffle : c'est le dernier soupir ou silence de mort... La soif d'air qui fait crier le mourant est celle-là même qui fait crier le nouveau-né : entre ces deux cris d'être, il y a le temps de la vie, le parcours d'une conscience, la trame d'un destin. Faire entendre sa voix, babiller, parler, chanter, rire ou pleurer, c'est vivre en sujet dans le monde des hommes »[17].

La voix commence avec le cri du nouveau-né, pour n'articuler des mots que plus tard dans l'existence. Un enfant en bas âge communique ses états émotionnels — ses besoins, affections et afflictions — avec des sons dont la tonalité a pourtant un sens, perçu ou non par ses parents. Cet exemple, commun à tous et ancré dans l'existence, illustre l'origine sensible de la voix. Cette sensibilité, nous en saisissons l'importance lorsque l'on pense au timbre, à la tonalité, au rythme d'une voix, sorte de communication para-verbale qui entoure le mot de l'adulte ou le cri de l'enfant. Le son de la voix — symbolisé par le cri, mode d'expression non articulé — porte en lui l'empreinte du corps. Son écho résonne en nous, instaurant une forme de dialectique entre voix intérieure et extériorité physique. Le son tire sa source de la chair et est en cela l'expression de ses mouvements — sa référence est la sensibilité, éminemment corporelle, qu'elle traduit en un langage plus ou moins articulé :

---

16 *Trésor de la Langue Française Informatisé*, définition de « voix ».

17 Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.11.

> « Les références de la sensibilité esthétique, chez l'homme, prennent leur source dans la sensibilité viscérale et musculaire profonde, dans la densité dermique, dans les sens olfacto-gustatifs, auditifs et visuels, enfin dans l'image intellectuelle, reflet symbolique de l'ensemble des tissus de sensibilité »[18].

L'origine sensible du matériau vocal se perçoit concrètement de plusieurs manières, qu'il nous faut détailler : l'antériorité de l'oralité sur l'écriture dans l'histoire humaine, qui trouve son explication dans l'évolution biologique de l'espèce (a) ; la charge de sens présente dans le bruit, réduit idéologiquement à un caractère de la source sonore dans le langage (b) ; la musicalité de la voix, qui consiste en un agencement de sons produisant un sens amoindri lorsque articulé dans le langage (c).

a) L'antériorité de l'oralité sur l'écriture dans l'histoire humaine.

Elle trouve son explication dans son évolution biologique. L'usage de la voix pour articuler un langage humain provient d'un développement de l'ouïe et de la structure musculaire qui permet l'émission d'un son. Les premières formes de vies étaient aquatiques. Ce n'est que lorsqu'une partie d'entre elles sont sorties de l'eau que des mammifères ont vu le jour, modifiant leur appareil auditif : « l'appareil des ébranlements liquides va se transformer en appareil de réception des ébranlements aériens », y introduisant les facultés de distinguer des timbres et des tonalités[19]. Grâce à cette amélioration, l'appareil vocal de l'homme s'est à la fois musclé et affiné. Nombre de chercheurs s'accordent ainsi sur le fait que c'est sa capacité vocale, qui provient d'une « musculature et une innervation musculaire très riche au niveau des joues, des lèvres et de la langue »[20], qui a progressivement distingué l'homme du singe : « Ce qui manque au chimpanzé, c'est non l'aptitude cérébrale, mais l'aptitude glottique »[21].

---

[18] Leroi-Gourhan, A., *Le Geste et la Parole*, éd. Albin Michel, tome 2, 1965, p.83.

[19] Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.14.

[20] ibid., p.17

[21] Morin, E., *Le Paradigme Perdu : la Nature Humaine*, éd. Seuil, 1973, p.53, in Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.17.

b) La charge de sens présente dans les bruits.

Cette dimension est discutée et théorisée par Christian Metz, sémiologue du cinéma, dans un passage de *Le Perçu et le Nommé* traitant des « objets sonores »[22]. Sa réflexion part d'un constat simple, prenant comme objet le mot *clapotis.* Nous attribuons spontanément cette désignation à des sons provenant de sources très diverses, ne pouvant nous mener qu'à la conclusion que « le clapotis existe comme un objet sonore autonome, avec les traits pertinents de son signifiant acoustique qui correspondent à ceux du signifié linguistique »[23]. Si les mots ne sont pas les choses elles-mêmes, il semble en effet que leurs sonorités présentent une analogie à ce qu'ils désignent. Le mot clapotis se définit comme un « bruit léger, répété et prolongé que provoque soit l'agitation de la surface de l'eau soulevée par des ondes de vagues courtes qui s'entrechoquent, soit l'eau calme qui se heurte régulièrement à un obstacle (bateau) ou aux bords (jetée, port) »[24]. Si cette définition identifie elle-même une variété de sources pouvant produire le clapotis, il semble qu'elle soit illustrée par la sonorité de son mot. Christian Metz distingue quatre ressemblances, tout en mentionnant qu'il en existe probablement d'autres, que nous paraphrasons ci-dessous[25] :

- Le bruit du clapotis est « faible ». Il s'oppose par exemple au « vacarme », au « hurlement » ou au « fracas », et retranscrit le lent et doux mouvement de l'eau dans le contexte qu'il décrit.
- Il est « discontinu », comme les vagues qui s'entrechoquent.
- Il est « double », car il se compose de deux sons successifs : le *c-l-*, à la valeur presque onomatopéique marqués plus longuement à l'oral, et le *-apotis*, qui se prononce en un mouvement. Cela n'est pas sans rappeler le léger heurt de l'eau mentionné dans notre définition.
- Le bruit généré par le mot « clapotis » est « liquide », attestant de la proximité de la sonorité du mot avec la chose qu'il décrit. A titre de comparaison, Christian Metz trouve dans les mots « frottement » et « raclement » un trait « solide », et dans les mots « chuintement » et « sifflement » un trait « gazeux ».

---

[22] Metz, C., « Le Perçu et le Nommé », *Essais Sémiotiques*, éd. Klincksieck, 1977, pp. 129-163.

[23] ibid., p.154.

[24] *Trésor de la Langue Française Informatisé*, définition de « clapotis ».

[25] Metz, C., « Le Perçu et le Nommé », *Essais Sémiotiques*, éd. Klincksieck, 1977, p.154.

La sonorité du mot et de l'objet sonore qu'il décrit coïncident, attestant de l'existence des sons comme des vecteurs de sens à part entière. Une sonorité n'est cependant jamais identifiée seule. Il est toujours préférable de qualifier sa source pour qu'elle soit tout à fait intelligible. Un « grondement » n'existe pas en lui-même, il faut ajouter qu'il s'agit du grondement « du tonnerre ». Cela, pour Metz, est arbitraire, et relève d'une idéologie qui privilégie la vision à l'ouïe car l'objet visuel est immédiatement identifiable, contrairement au son, dont les contours restent flous : « nous considérons comme évident qu'un étendard est un objet (tout court), mais pour un hululement nous hésitons : c'est un infra-objet, un objet seulement sonore »[26]. L'origine sensible du matériau sonore en complique l'identification mais, comme l'exemple du « clapotis » semble le souligner, est également antérieure au langage en tant qu'articulation de mots puisque les sonorités de ces derniers sont en partie construits en analogie avec la source qu'ils désignent.

c) La musicalité de la voix.

La dimension musicale de la voix devient, grâce à l'exemple que nous venons de développer, progressivement plus évidente. Elle ne s'arrête pourtant pas aux « objets sonores », aux mots qui les désignent et à leurs sonorités. Elle est perceptible dans ce que Roland Barthes nomme le « grain de la voix », soit « un mixte érotique de timbre et de langage, et peut donc être lui aussi, à l'égal de la diction, la matière d'un art : l'art de conduire son corps ». La création de ce concept par le sémiologue trouve son origine en la nécessité de lier la sonorité de la voix à son ancrage dans le corps, thèse qu'il développe en dissertant sur la quête d'une « écriture à haute voix » qui en emprunterait les qualités :

> « (…) ce qu'elle [l'écriture à haute voix] cherche (dans une perspective de jouissance), ce sont les incidents pulsionnels, c'est le langage tapissé de peau, un texte où l'on puisse entendre le grain du gosier, la patine des consonnes, la volupté des voyelles, toute une stéréophonie de la chair profonde : l'articulation du corps, de la langue, non celle du sens, du langage »[27].

Un grain sonore, et *a fortiori* musical — un « art de la mélodie »[28] puisque consistant en l'agencement de ces sons dans le langage en préservant leur qualité corporelle. Sans réduire le

---

[26] ibid., p.156.

[27] Barthes, R., *Le Plaisir du Texte*, éd. Seuil, 1973, pp.104-105.

[28] ibid.

concept de Barthes, il est possible d'en appréhender quelques composantes en qualifiant la voix en termes musicaux. Marie-France Castarède distingue trois paramètres dans la voix, en tant que « support acoustique de la parole » : la fréquence, l'intensité et le timbre[29] :

- *La fréquence* désigne la tonalité plus ou moins haute de la voix, qui oscille autour d'une tonalité moyenne appelée « fondamental usuel de la parole ». Le degré de l'oscillation est propre à chaque voix, et dépend de facteurs « physiologiques, familiaux et sociaux ».
- *L'intensité* peut être traduite vulgairement en volume de la voix. Celui-ci a également une signification, notamment psychologique pour Castarède, qui cite en exemple la différence d'intensité entre une voix timide et celle d'une personne affirmée.
- *Le timbre* se divise entre un « timbre vocalique », qui permet d'identifier les voyelles, et un « timbre extra-vocalique », la sonorité de la voix d'un individu, qui est selon Castarède attachée à la personnalité — « une émanation de la personnalité profonde du sujet »[30].

Les composantes musicales de la voix sont la trace du sensible dans nos modes d'expression, et sont en cela l'émanation de l'individu qui les prononce. L'art musical est ainsi celui de la transcription immédiate du sensible qui origine l'existence, loué par une multitude d'auteurs. Pour exemple, Schopenhauer voit en la musique l'objectivation du monde en tant que volonté et représentation — un langage des essences que le langage articulé ne fait qu'effleurer : *[La musique n'est par conséquent en aucun cas comme les autres arts, une copie des Idées, mais une copie de la Volonté elle-même, l'objectivité dont sont faites les Idées. Pour cette raison, l'effet de la musique est bien plus fort et pénétrant que celui des autres arts, qui ne parlent que de l'ombre quand la musique traite de l'essence]* :

> « Therefore music is by no means like the other arts, namely a copy of the Ideas, but a copy of the will itself, the objectivity of which are the Ideas. For this reason, the effect of music is so very much more powerful and penetrating than is that of the other arts, for these others speak only of the shadow, but music of the essence »[31].

---

[29] Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.26.

[30] ibid., p.28.

[31] Schopenhauer, A., *The World as Will and Representation*, traduit par E.F.J. Payne, Volume I, éd. Dover, 1966, p.257, in Storr, A., *Music and the Mind*, éd. Harper Collins, 1993, p.140.

Son lien étroit avec un état sensible de l'expression introduit la musique comme origine de la langue — « La musique prévient tout énoncé, puisque le langage et les concepts se nichent dans l'irruption du flux acoustique, c'est-à-dire viennent après »[32].

Si le langage conserve cette qualité musicale, son articulation d'un lien arbitraire entre un signifiant et un signifié[33] en réduit à la fois le sens et la portée. Rousseau interprète cela comme la descente de la langue dans le langage, du sensible dans l'intelligible, du *mythos* dans le *logos*, qui s'inscrit dans l'histoire des techniques, qui font société, et supposent d'user du langage à des fins pratiques :

> « On ne commence pas par raisonner mais par sentir... L'origine des langues n'est point due aux premiers besoins des hommes... D'où peut donc venir cette origine ? Des besoins moraux, des passions. Toutes les passions rapprochent les hommes que la nécessité de chercher à vivre force à se fuir. Ce n'est ni la faim, ni la soif mais l'amour, la haine, la pitié, la colère qui leur ont arraché les premières voix... Voilà pourquoi les premières langues furent chantantes et passionnées avant d'être simples et méthodiques »[34].

2. La voix, support et médiation d'un arbitraire du signe.

La voix semble consister en la conjugaison de qualités musicales aux référents sensibles et d'un signe linguistique qui dénote arbitrairement une chose. Mais elle est aussi la tentative de son dépassement, de la réduction de l'« écart » entre une essence des idées et la parole humaine[35]. Nous pourrions résumer cela en un effort de connotation, effectué en augmentant la dimension musicale du langage, et ce par différents biais : le chant et la poésie. Le chant, en ce qu'il incorpore la musique dans le langage, est l'exemple qui fait le plus évidemment suite à nos précédentes considérations. La poésie, parlée ou écrite, consiste pourtant en un même effort à la fois métaphorique et musical d'ajouter une couche de sens au langage. Le poème des « Voyelles » de Rimbaud, qui tisse une relation synesthésique entre des lettres et des images métaphoriques tout en

---

[32] Serres, M., *Esthétiques. Sur Carpaccio*, cité par Charles, D., *Le Temps de la Voix*, éd. Delarge, 1978, p.153.

[33] Barthes, R., *Éléments de Sémiologie*, in *Communications*, 4, 1964, « Recherches Sémiologiques », pp.91-135. Disponible à l'adresse suivante : https://www.persee.fr/doc/comm_0588-8018_1964_num_4_1_1029

[34] Rousseau, J.-J., *Essai sur l'Origine des Langues*, éd. Aubier, 1973, pp.95-96.

[35] Genette, G., « Langage Poétique, Poétique du Langage », in *Figures II*, éd. Seuil, 1979, pp.123-153.

usant d'un schéma commun à nombre de poèmes, fait de rimes et de rythmes, en est une célèbre illustration.

Nous souhaitons cependant nous concentrer sur le langage parlé, qui est le plus récurrent en contexte radiophonique. Dans ce cadre aussi, la modulation des sons de la voix lors de l'élocution d'un énoncé est porteuse de connotation. Iván Fónagy, chercheur en psycho-phonétique, voit dans la notion de « style vocal » — soit une « manière expressive de prononcer »[36] — la possibilité d'une modulation du message linguistique porteuse de sens. L'énonciation d'un message vocal à un destinataire consisterait en un double encodage : un premier, linguistique, puis un second, qui reproduit un mouvement de l'âme, une émotion ressentie ou que l'on souhaite véhiculer dans la sonorité de l'énoncé. Le destinataire doit décoder ces deux couches de sens, la modulation sonore pouvant altérer la stricte signification linguistique, qu'il doit retrouver. Iván Fónagy présente sa théorie en ces termes[37] :

> « Il faudrait admettre, par conséquent, deux actes successifs d'encodage : un encodage linguistique qui transforme un message global, une idée, en une séquence de phonèmes, et un deuxième codage — qui coïncide admirablement avec l'acte de mise en sons des phonèmes — au cours duquel le message *secondaire*, gestuel, est greffé sur le message primaire ».

L'auteur donne un exemple. Lorsqu'une mère s'adresse à son enfant en modulant la phrase « Mais si, voyons » en ce qui donnerait phonétiquement « Mais *su*, voyons », remplaçant le son du *i* par un *u*, elle ajoute à un premier encodage grammatical un second encodage de nature sonore, qui ici pourrait véhiculer « la moue affectueuse préfigurant le baiser »[38].

Deux points sont à préciser à propos de ce second encodage. Premièrement, il s'agit d'un processus para-linguistique qui se situe « (…) sur un autre niveau mental. Le message échappe à la conscience : on ne perçoit que son effet »[39]. Cet élément résonne avec le caractère inexplicable car originaire de l'information sensible que nous avons mentionné précédemment. Deuxièmement, ce second encodage, bien qu'il use du son en tant que matériau sensible, n'est pas une faculté naturelle de l'homme. Il consiste davantage en une coutume socialisée, et par là-même « naturalisée »[40], dans

---

36 Fónagy, I., *La Vive Voix. Essais de Psycho-phonétique*, éd. Payot, 1983, p.12.

37 ibid., p.14.

38 ibid., p.13.

39 ibid., p.16.

40 Barthes, R., *Mythologies*, éd. Seuil, 1957.

le langage. L'usage du son dans la voix demeure ainsi contraint par l'arbitraire du signe. Il concourt à la communication d'un état dans une situation de communication d'abord instaurée par la grammaire de chaque interlocuteur.

3. La *vive voix*, incarnation de la personnalité.

Notre exploration des propriétés de la voix et de son statut au sein du langage nous amène au constat suivant : elle conjugue une émotion (l'expression d'un sensible musical) et la codification de la grammaire (arbitraire du signe qui en réduit le sens). La voix contient une part non-intelligible, véhicule d'un sens aux modalités différentes, que l'on perçoit, comprend, et s'approprie comme la mise en contexte d'une interaction verbale ou d'une situation d'écoute. Elle est un support du signe linguistique, qui contraint son expression ; elle est sa médiation, car elle lui ajoute une connotation d'autant plus prononcée lorsqu'elle est « vive », pour reprendre les termes d'Iván Fónagy. La voix transmet des émotions à l'écho universel, à la manière de la musique, mais que l'individu qui la prononce s'approprie. Elle est une expression personnelle, individuelle, dont le timbre, que nous avons décrit précédemment est une composante. Lorsque la vive voix est transposée telle quelle à l'écrit, une déperdition se fait de suite remarquer. Iván Fónagy le formule ainsi :

> « La *vive voix* s'oppose en Français, comme en d'autres langues, à la *lettre morte*. Les métaphores ne sont jamais gratuites. Si elles ont l'air de l'être, ceci veut simplement dire qu'elles vont au-delà de la pensée consciente et expriment quelque chose que nous savons sans le savoir. L'opposition de ces deux termes indique que la vivacité est inhérente à toute communication vocale »[41].

Le « grain de la voix » que Barthes fantasme de voir traduit à l'écrit provient des émotions et de la sensibilité du locuteur, que l'écrit ne peut froidement répliquer. L'usage d'une métaphore mortuaire — la « lettre morte » chez Fónagy ou la « toilette du mort »[42] que Barthes utilise pour décrire l'effet de la transcription sur des paroles — l'illustre avec cohérence. L'écriture n'emploie non plus un *style vocal* mais un *style littéraire*, qui contient plusieurs registres de langue et de discours. Ce style vise également à conjurer l'arbitraire du langage dans une perspective poétique de réduction d'un « écart » que nous avons décrit.

---

41 Fónagy, I., *La Vive Voix. Essais de Psycho-phonétique*, éd. Payot, 1983, p.9.

42 Barthes, R., *Le Grain de la Voix. Entretiens 1962-1980*, éd. Seuil, 1981, p.9.

Cette conception ne fait cependant pas légion. Jacques Derrida va jusqu'à renverser la structure métaphysique que nous employons depuis le début de notre réflexion — où la sonorité découlerait d'une sensibilité originelle et originaire du langage —, voyant en l'écriture la *trace* d'un méta-code linguistique régissant le signe oral et écrit sans que celui-ci ne réponde à une stricte relation d'un signifiant à un signifié, car l'interprétation d'un signifiant est plurielle[43]. Cela n'est pas sans rappeler l'analyse de l'« objet sonore » par Christian Metz, dont le sens est multiple (le clapotis de l'eau peut correspondre à plusieurs situations, affirmant son existence autonome en tant que signe). Cette dernière s'ancre néanmoins dans le cadre du structuralisme instauré par la convergence de plusieurs disciplines au XX^{e} siècle (dont la linguistique saussurienne) que Derrida, en tant que post-structuraliste, souhaite déconstruire.

A la lumière de ces contradictions qui traversent la linguistique, nous ne pouvons nous prononcer sur la validité du principe d'antériorité de l'une des composantes de la voix sur une autre. Nous remarquons du moins que la voix est le théâtre d'une tension entre les sonorités des émotions d'un individu et l'arbitraire d'un langage qui fait société. C'est là sa nature profonde. La voix est *vive*, elle est incarnée par une personnalité qui l'énonce, à l'oral comme à l'écrit, et fait résonner des émotions intérieures dans le monde, chez les autres. La voix se réalise dans le souffle du corps qui la projette. Sa nature est celle du sentiment intérieur, qu'elle extériorise :

> « La voix exprime premièrement l'invisible, l'intime, le sentiment. Le son est frère de l'âme, dit Hegel. Le pur art vocal ne décrit rien que les vicissitudes de l'émotion totale. Le son est de volonté. Il surmonte la nature. Le chant est une lutte contre le cri »[44].

Le lien entre la voix et le parcours de nos émotions, intime et invisible, que nous avons tenté d'expliciter, est selon nous au fondement des qualités du média radio en tant que dispositif acousmatique (qui ne laisse pas voir la source sonore). La radio provoque plusieurs effets (d'intimité, de communauté, d'évocation et de réel) sur ses auditeurs — que nous rassemblons sous le terme d'authenticité — qui proviennent de l'ancrage de son matériau vocal dans le sensible.

---

43 Derrida, J., *De la Grammatologie*, éd. de Minuit, 1967.

44 Alain, « Vingt Leçons sur les Beaux Arts », in *Les Arts et les Dieux*, éd. Gallimard, coll. La Pléiade, 1968, pp.505-506, cité par Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.60.

B. LA VOIX À L'ORIGINE DES EFFETS PROPRES AU MÉDIA RADIO : L'AUTHENTICITÉ D'UN DISPOSITIF ACOUSMATIQUE.

1. La voix, condition de possibilité des effets du média radio : l'authenticité en tant que reflet des sens.

La radio transporte la voix d'un ou de plusieurs locuteurs vers une communauté d'écoute, historiquement au travers de la bande FM. Sur le numérique s'ajoute la radio IP, qui permet à une radio d'émettre au travers de Réseaux Privés Virtuels (VPN), et des webradios. A l'écoute de son poste, l'auditeur interagit — de manière passive, car limitée à l'écoute — avec un contenu parlé et sonorisé — une émission de débats, un programme musical, ou encore une libre antenne. De cette relation émerge une sensation de connexion émotionnelle à la voix entendue, qui d'après ce que nous avons étudié jusqu'à présent provient de son référent sensible, et que nous pourrions résumer en un sentiment d'authenticité. Ce dernier rassemble selon nous plusieurs effets produits par l'écoute du média radio : d'intimité (a), de communauté (b), de réel (c), et d'imaginaire (d).

a) L'effet d'intimité : la voix de radio comme projection d'une intériorité.

La voix est une projection de notre intériorité. L'écoute de celle d'un autre fait écho à notre voix intérieure. La voix intérieure correspond à notre conscience — le *cogito ergo sum* de Descartes, « je pense donc je suis » que l'on s'adresse à soi-même, en est l'illustration —, qui dans le moment où le sujet l'extériorise se voit déniée car expulsée hors du corps. La voix intérieure est le terrain qui glisse dès qu'elle devient langage. Elle est éminemment liée à la vie du sujet, vive et incarnée. Avant la radio, l'invention du phonographe au XIX^{e} siècle a pour cette raison fasciné les publics. La voix pouvait désormais émaner d'une machine à n'importe quel moment, elle n'était plus liée au présent existentiel de son locuteur : *[« En écoutant l'enregistrement, on pensait que l'auditeur serait transporté dans le lieu et le moment de l'événement d'origine. Il/elle expérimenterait la voix reproduite dans toute sa présence — presque comme si elle était là, au milieu du salon »]* :

> « In listening to the recording, the auditor, it was assumed, would be transported to the time and place of the original event. He/she would experience the reproduced voice in all its presence — almost as if it were there, in the middle of the living room »[45].

La radio, qui consiste en la mise en ondes de ce processus dans un flux en direct, produit le même type de fascination, car elle fait résonner l'intériorité de l'individu qui l'écoute. Gaston Bachelard, philosophe des sciences, exprime ainsi son attrait pour la radio, support de rêverie :

> « La radio est vraiment la réalisation intégrale, la réalisation quotidienne de la psyché humaine. Le problème qui se pose à cet égard n'est pas purement et simplement un problème de communication ; ce n'est pas simplement un problème d'information ; mais, d'une manière quotidienne, dans les nécessités non seulement d'information, mais de valeur humaine, la radio est chargée de présenter ce qui est la psyché humaine »[46].

Elle produit une intimité car son écoute résonne avec l'intériorité de l'auditeur, qui a l'impression qu'on ne s'adresse qu'à lui.

b) Un effet de communauté permis par la diffusion en direct.

La radio double l'effet d'intimité qu'elle partage avec le phonographe en s'adressant à une communauté en direct. Les publics à l'écoute entendent tous le même message, au même instant. La voix de radio est portée à une multiplicité de paires d'oreilles, auxquelles elle s'adresse personnellement en simultanéité. L'intimité du média radio est ainsi renforcée par la fédération d'une communauté d'auditeurs, à l'échelle d'une nation avant la montée en puissance de la télévision, qui lui a progressivement fait de l'ombre dans les usages quotidiens. Les communautés d'auditeurs désignent désormais davantage le public d'une station, ou d'un programme, suscitant un sentiment d'appartenance fort, pour exemple caractéristique des jeunes publics à l'écoute de libres antennes, notamment sur les stations musicales en soirée. Les personnes se socialisent avec la radio. En tant que média de masse, la radio remplit une fonction similaire à celle de la « télévision

---

45 Citation traduite de l'Anglais, in Dyson, F., *Sounding New Media. Immersion and Embodiment in the Arts and Culture*, éd. University of California Press, 2009, p.46.

46 Bachelard, G., « Rêverie et Radio », in *Droit de Rêver*, éd. PUF, 1970, p.216.

cérémonielle », théorisée par Dayan et Katz[47]. Elle crée l'événement, et est un sujet de discussion collective. Marshall McLuhan décrit cet aller-retour entre l'individu et la société que produit la radio dans son ouvrage *Pour Comprendre les Médias* :

> « La radio touche les gens dans leur intimité. C'est une relation de personne à personne, qui ouvre tout un monde de communication entre l'auteur-speaker et l'auditeur. C'est là le côté direct de la radio. C'est une expérience privée. Des profondeurs subliminales de la radio surgit l'écho résonnant des trompes tribales et des tambours antiques. C'est là quelque chose d'inhérent à la nature même de ce médium, qui a le pouvoir de transformer l'individu et la société en une seule et même chambre réverbérante »[48].

Le transport de la voix dans cette « chambre réverbérante » que constitue la radio était même considéré comme l'avènement d'une « seconde oralité » par Walter Ong, un linguiste américain, car elle décuple la portée millénaire d'une fonction de la communication orale — faire communauté : *[Comme l'oralité première, la seconde oralité a généré un fort sentiment de communauté, mais incommensurablement plus grand que celui de la culture orale primaire]* « Secondary orality has generated a strong group sense, but [it] generates a sense for groups immeasurably larger than those of primary oral culture »[49]. Ce sentiment de communauté, effectif à l'échelle d'une nation lorsque la radio était moins concurrencée par la télévision, est décuplé car il est forgé par la propagation en direct à un large public de contenus d'information et de divertissement dans une époque marquée par les angoisses de l'industrialisation. La radio ancre la parole d'un individu, le *speaker* de radio, et son écoute par une communauté d'auditeurs dans le monde.

Il semble ainsi que la portée sociale de la radio dépende fortement d'une notion de rendez-vous, appuyée par une grille de programmes venant structurer un flux en direct, créant ainsi des habitudes chez ses auditeurs. L'instantanéité de la radio ancre sa parole et son écoute dans le temps présent. Guillaume Duhamel, concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, traduit un effet similaire en évoquant la radio de son adolescence :

---

47 Dayan, D., Katz, E., *La Télévision Cérémonielle. Anthropologie et Histoire en Direct*, éd. PUF, 1996.

48 McLuhan, M., *Pour Comprendre les Médias*, éd. Seuil, 1968, pp.341-342.

49 Citation extraite de Ong, W., *Orality and Literacy*, éd. Methuen, 1982, p. 136, rapportée par Dyson, F., *Sounding New Media. Immersion and Embodiment in the Arts and Culture*, éd. University of California Press, 2009, p.50.

> « Plus largement, la radio a historiquement une fonction sociale. Imagine des gens dans les années trente qui écoutent les informations ensemble. Puis elle a été remplacée par la télé. Elle a eu un rôle de sociabilisation. Une émission de radio devenait un événement, et les gens se donnaient rendez-vous pour écouter une émission de radio. Même plus récemment. Quand j'étais au collège, on savait que tel soir il allait se passer ça dans une émission. On était impatients, chacun était chez soi à l'écoute. Et le lendemain dans la cour de récré on parlait de l'émission de la veille. Il y avait un aspect social et de communauté »[50] .

c) L'effet de réel : l'écoute de la voix dans un dispositif acousmatique.

L'ancrage de la radio dans le temps de son énonciation est le premier élément constitutif de ce nous pourrions qualifier d' « effet de réel ». Ce concept est développé par Roland Barthes pour qualifier le sentiment de réalité que provoque la description de différents détails dans *Madame Bovary* de Flaubert. Ces détails, presque inutiles, ancrent le récit dans une réalité existentielle car l'authenticité de leur description signifie, de manière fictive, l'existence de « référents réels »[51]. Le grain de la voix, autre concept de Roland Barthes que nous avons décrit plus tôt, remplit une fonction similaire. Les détails de l'élocution de la voix de radio — ses respirations, hésitations, accentuations et exclamations — manifestent l'existence d'un référent sensible. Le contexte de l'énonciation radiophonique restitue un *grain* qui en exprime la réalité, allant parfois jusqu'à saisir l'écoute, la subjuguer. L'usage de l'adjectif « poignant » par Alban Rivoal, chargé de production à Fun Radio, alors qu'il tentait de qualifier ce que l'écoute du témoignage d'un policier étant intervenu au Bataclan lors des attaques à Paris en 2015 avait provoqué en lui, traduit cet aspect :

> « Un exemple très poignant : suite aux attentats du Bataclan, le commissaire qui est rentré le premier dans la salle a fait un témoignage sur plusieurs radios dont RTL. C'est poignant, parce que la personne explique, et avec le ton de sa voix... Tu imagines tout. Il explique qu'il prend son téléphone, qu'il dit "adieu" à sa famille parce qu'il ne sait pas ce qui va lui arriver, et il rentre dans le lieu. Et ça continue comme ça. Ça te glace le sang. Tu passes par plein d'émotions rien qu'avec la voix. Je pense qu'on peut faire la même chose en télé, mais rien qu'à la voix, ce témoignage, en audio, était poignant et je n'avais pas besoin d'avoir d'images »[52].

---

[50] Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

[51] Barthes, R., « L'Effet de Réel », in *Le Bruissement de la Langue*, éd. Seuil, 1984, pp.179-187.

[52] Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp.154-158.

Cet épithète pourrait s'interpréter comme la capacité de la voix à construire un *punctum*, concept utilisé par Roland Barthes pour décrire les qualités sensibles de la photographie : « (…) car *punctum*, c'est aussi : piqûre, petit trou, petite tache, petite coupure — et aussi coup de dés. Le *punctum* d'une photo, c'est ce hasard qui, en elle, *me point* (mais aussi me meurtrit, me poigne) »[53]. Il y a dans la voix une esthétique du sensible au travers des détails de ses sonorités, de son *grain*, qui a un effet de réel sur les auditeurs.

d) La voix de radio, médiation de l'imaginaire.

Intime, créatrice de lien social, et ancrée dans une réalité — un présent de l'écoute —, la radio a sur ses auditeurs un pouvoir d'évocation. L'écoute du seul son de la voix fonctionne par métonymie, « un terme qui, en langue, en signifie un autre, celui-ci étant, au départ, associé à la première par un rapport de contiguïté »[54], voire de synecdoque, « figure de rhétorique procédant par extension ou restriction de sens d'un terme »[55], puisque le fait de ne recevoir que le son suppose que l'auditeur recrée mentalement la situation dans laquelle il est plongé. La voix devient ainsi le matériau même de l'élaboration du récit dans un dispositif dit « acousmatique », « c'est-à-dire qu'on l'entend sans en voir la source »[56], qui a la propriété de mobiliser l'imaginaire de celui qui l'écoute. Patrick Charaudeau décrit une radio qui joue avec « les ressources de l'oralité et du son, pour représenter un monde d'événements dont la présence se fait par le biais d'évocations produites par le son et la voix »[57]. Pierre Laurent, chef de produit marketing chargé de la création et de l'analyse sonore au groupe radio RTL, évoque ce lien entre radio et imaginaire :

> « Sans l'image, tout devient plus grand, tout devient plus personnel. Chacun imagine ce qu'il écoute. C'est un peu comme la lecture. Quand tu lis un roman, c'est toi qui vas te projeter dans l'histoire, imaginer le visage du héros, les corps,

---

53 Barthes, R., *La Chambre Claire. Note sur la Photographie*, éd. Seuil, 1980, p.49.

54 *Trésor de la Langue Française Informatisé*, définition de « métonymie ».

55 *Trésor de la Langue Française Informatisé*, définition de « synecdoque ».

56 *INA Global*, Smati, N., Fievet, A.-C., « A la radio, la voix donne à écouter et à voir », 04/12/2017. Disponible à l'adresse suivante : https://www.inaglobal.fr/radio/article/la-radio-la-voix-donne-ecouter-et-voir-10010

57 Charaudeau, P., « Le contrat de communication de l'information médiatique », in Lancien, T., *Médias : faits et effets*, Paris, Hachette, 1994, p.14, cité par Smati, N., Fievet, A.-C., *INA Global*, « A la radio, la voix donne à écouter et à voir », 04/12/2017. Disponible à l'adresse suivante : https://www.inaglobal.fr/radio/article/la-radio-la-voix-donne-ecouter-et-voir-10010

> les paysages... On dit que le son stimule à plein tubes notre imagination et c'est vrai, parce qu'il n'y a pas de limites, là où l'image t'impose quelque chose. C'est toi qui fais ton image mentale. Donc le son c'est un stimuli qui va engendrer en chacun de nous une réaction propre »[58].

Guillaume Duhamel y voit même la singularité du média radio : « Ce sont tes oreilles qui ont des yeux. Pendant longtemps, en radio, tu ne savais pas qui était derrière le micro. Tu entendais une voix et ton imaginaire essayait de te représenter un visage »[59].

La radio, parce qu'elle travaille et focalise l'attention sur la voix, construit une relation d'authenticité avec ses auditeurs. Ce terme semble rassembler les effets que avons décrit : l'intimité qu'instaure l'écho de voix projetées et intérieures, la fédération d'une communauté d'écoute, l'effet de réel que suscite l'ancrage dans un présent de l'énonciation que permet la diffusion en direct, et le pouvoir d'évocation accentué par un dispositif acousmatique. La radio ancre la parole dans le présent, dans le *hic et nunc* (l'ici et maintenant) de la relation d'un locuteur et d'un auditeur qu'elle porte avec une voix qui noue ces deux corps. Elle restaure ce que Walter Benjamin déclarait perdu avec l'avènement d'une possibilité de reproduction technique des œuvres d'art, un *hic et nunc* qu'il qualifiait lui-même d' « authenticité », se distinguant en cela du phonographe :

> « L'authenticité d'une chose réside dans tout ce qu'elle peut transmettre d'elle depuis son origine, de sa durée matérielle à son pouvoir d'évocation historique. Puisque celui-ci se fonde sur celle-là, si la chose tombe dans la reproduction, là où sa durée matérielle s'est dérobée aux hommes, son pouvoir de témoignage historique s'en trouve tout aussi ébranlé. S'il est vrai que cela n'est que cela, ce qui n'en est pas moins ébranlé, c'est l'autorité même de la chose »[60].

Nous avons décrit les qualités du média radio, qui selon nous proviennent des modalités de sa médiation de la voix. Le travail de cette dernière entre pourtant en jeu avant sa réception acousmatique par l'auditeur, dès son énonciation dans le micro.

---

58 Entretien avec Pierre Laurent, Chef de produit chargé de la création et de l'analyse sonore au groupe radio RTL, mené par Clément Mellouet. Annexe 12, pp.187-191.

59 Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

60 Benjamin, W., *L'Oeuvre d'Art à l'Époque de sa Reproductibilité Technique*, 1955, éd. Allia, 2016, p.21.

2. La voix de radio, travail de l'authenticité comme une relation de proximité.

Au-delà des effets qu'elle produit, l'identité de la radio est attachée à cette notion d'authenticité, qu'elle travaille dans le ton de ses interventions. Ce média remplit une fonction d'accompagnement des auditeurs, et ce de deux manières : elle crée des habitudes d'écoute, structurées par une grille des programmes proposant des rendez-vous, et elle est un média du quotidien, compagnon de ses auditeurs dans leurs tâches journalières car ne nécessitant que l'ouïe, contrairement à la télévision qui mobilise la vue, ou la presse, qui convoque également le toucher. La connexion qu'instaure le référent sensible de la voix s'adjoint aux situations d'écoute du média radio, qui imprègne la vie de ses auditeurs. Le *prime* en radio est ainsi le matin — à l'inverse de la télévision, qui voit ses audiences augmenter le soir — car elle accompagne la réalisation d'activités quotidiennes — quand on se prépare à partir au travail, ou lorsqu'on s'y rend en voiture. Le ton de la voix et la fonction d'accompagnement du média radio sont fondateurs du rapport qu'ont les auditeurs avec ce média. Ces deux notions sont prégnantes dans la description que Sacha Ephraim, régisseur-monteur chez RTL, effectue de sa relation à la radio :

> « La radio c'est hyper subjectif, hyper personnel. C'est un média beaucoup plus intime que la télé. Pendant que tu écoutes la radio tu vas te brosser les dents, tu es dans ta douche, tu es dans ta voiture... C'est des moments où tu fais d'autres choses, mais il y a ce petit truc qui t'accompagne tout le temps. Du coup les voix de la radio, que ce soit il y a dix ans, il y a trente ans ou aujourd'hui, elles t'accompagnent. Elles sont parfois rassurantes, tu lies quelque chose avec elles, même si elles ne te connaissent pas. Une relation particulière se crée. Moi tous les matins, quand je vais au boulot et que je commence tôt, je suis avec Jean-Jacques Bourdin sur RMC parce qu'il a une voix hyper posée, hyper grave. Je ne suis souvent pas d'accord avec ce qu'il dit, mais il y a un petit quelque chose... Jean-Jacques Bourdin, c'est mon truc du matin. Parce qu'il a une voix rassurante. Quand ça ne va pas, ce mec là, même s'il ne parle pas du tout de moi, ça me rassure de me dire qu'il est toujours là »[61].

Cette relation de proximité, les animateurs de radio l'incarnent avec leurs voix dans ce que l'on pourrait nommer leurs « phonostyles ». Ce terme désigne « le style singulier que chacun peut se donner, se façonner », en « lien avec la construction de la personnalité », « vu comme un ensemble d'ajustements à une situation », dans la manière de parler[62]. Lorsqu'il prend le micro,

---

[61] Entretien avec Sacha Ephraim, Régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 6, pp. 164-167.

[62] Boula de Mareüil, P., « Qu'est ce qu'un phonostyle ? », in *Nouveaux Cahiers de Linguistique Française*, 2014, 31, pp. 9-19. Disponible à l'adresse suivante : http://clf.unige.ch/files/9014/4102/7363/1_Boula_9-19.pdf

l'animateur radio incarne une *persona* radiophonique. Plus précisément, si la voix est une expression de notre personnalité dans la vie courante, elle représente une *persona* en contexte radiophonique. Ce terme, qui provient du Latin *personare* — « parler au travers » — désignait le masque que portaient les comédiens d'une pièce de théâtre. Il se rapproche de la notion de « masque social » théorisée par Erving Goffman, sociologue américain, qu'il applique à « la mise en scène de la vie quotidienne ». Selon lui, nous interpréterions une multiplicité de rôles pour nous adapter aux situations diverses qu'engendre la sociabilité[63]. Si l'ouvrage de Goffman emprunte la voie de la sociologie, où la notion de phonostyle n'est qu'annexe, l'auteur était préoccupé par ce sujet. Il a ainsi consacré dans l'édition américaine de ses *Façons de Parler* un chapitre au « *radio talk* », non disponible en Français[64]. Adoptant un point de vue réflexif sur le discours des animateurs radio, il identifie les fautes de prononciation et leur correction en direct par le locuteur comme révélatrices de la convention — du « contrat » pour reprendre les termes de Patrick Charaudeau ou d'Eliseo Veron — qu'elles viennent briser : *[L'animateur radio n'est censé nous apparaître que sous couvert de ce qui lui a été préparé, presque comme s'il devait coller au rôle qu'on lui a attribué dans une pièce de théâtre. Que l'on requière de lui une lecture à haute voix ou une parole spontanée, il est obligé de compresser ou d'étirer sa parole pour qu'elle dure exactement le temps qui lui a été dédié, remplissant l'espace entre les signaux «* on *» et «* off *». En vue de cet idéal, toute erreur notoire peut non seulement créer des associations incohérentes (voire mal nous informer), mais également dévier du flux obligatoire, révélant quelqu'un en train de trébucher — quelqu'un de maladroit — au lieu d'une personne programmée pour l'occasion]* :

> « [The radio speaker] is to appear to us only in the guise that his prepared material has planned for him, almost as though he were to hold himself to the character allotted to him in a play. And whether aloud reading or fresh talk is required of him, he is obliged to compress or stretch his talk so that it lasts exactly as long as the time allotted, just filling up the space between his 'on' and 'off' cues. Given this ideal, any noticeable fault may not only introduce irrelevant associations (if not misinforming us), but also divert the obligatory stream, presenting the view of someone stumbling — indeed a view of a stumbler — instead of a view of the person who has been programmed for the occasion »[65].

---

63 Goffman, E., *La Mise en Scène de la Vie Quotidienne. 1 : La Présentation de Soi*, éd. Minuit, 1973.

64 Goffman, E., « Radio Talk », in *Forms of Talk*, éd. University of Pennsylvania Press, 1981, pp.197-327.

65 ibid., p.320.

La radio instaure, comme tout média, un contrat avec ses auditeurs, dans lequel l'animateur, en ce qu'il parvient à le respecter, manifeste une compétence professionnelle. Toute erreur le décrédibilise et rompt la bonne conduite du contrat, que l'animateur incarne en son nom. Judith Clopeau et Madeleine Renouard, chercheuses à l'université de Londres, qualifient cette compétence incarnée comme la marque d'une « *homo vox* », rappelant le caractère central de la voix jusque dans les modalités de la communication radiophonique : « Son acte de parole sera ainsi justifié, homologué, légalisé, donc professionnel : « Homo Vox ». « Homo Vox » parce que les textes produits existent, éphémères, par la voix de quelqu'un qui a réussi à établir, en quelque sorte, un équilibre harmonieux entre l'Institution, l'Individu et les Indices d'écoute »[66]. S'il ne permet pas d'erreur, le contrat de communication radiophonique standardise les voix que l'on peut entendre à l'antenne : « les voix trop aiguës, trop basses, ou présentant une trop grande « non-conventionnalité », sont exclues. Hormis des interventions ponctuelles/saisonnières, pas d'enfants, d'adolescents, ni de personnes du troisième âge. Toujours peu d'accents régionaux (à l'état natif) et, quand ils sont présents, leur caractère iconique est plus fortement marqué que celui des autres voix »[67].

La voix de radio est donc contrainte par une pluralité de dimensions, établies à l'échelle d'une station par la direction des programmes, et mises en œuvre par l'animateur lorsqu'il parle. Ces dimensions sont d'abord la conséquence d'une représentation du public auquel le média et l'animateur souhaitent s'adresser, affinée par les mesures d'audiences actuelles qui raisonnent en termes de cibles — commerciales dans le cas des radios privées, ou plus généralement de population. Allan Bell, chercheur en sociolinguistique, parle en ce sens d' « *audience design* » (que l'on pourrait traduire par « conception des publics »), qui a une influence sur le ton du discours. Il établit une typologie des publics, allant du « destinataire » — que l'on connaît, dont l'existence est en cela certifiée, et auquel on s'adresse directement dans une situation d'interaction immédiate — à « l'indiscret », dont le locuteur ne connaît pas même l'existence. Au milieu, « l'auditeur », que l'on connaît et est en cela certifié, mais auquel on ne s'adresse pas directement, puisque la relation fait l'objet de la médiation du dispositif radiophonique[68]. L'animateur essaie d'employer un phonostyle qui minore la distance qui le sépare de ses auditeurs, instaurant une impression de proximité, mais qui est nécessairement généralisé à l'ensemble des auditeurs d'une station. Lorsque cette tension est

---

66 Clopeau, J., Renouard, M., « Quelques remarques sur la « communication » radiophonique », in Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984, p.23.

67 ibid., p.23.

68 Bell, A., « Language Style as Audience Design », in *Language in Society*, 13(2), pp.145-204.

mal dosée, la parole de l'animateur court le risque de la perte de sincérité, comme l'explique Guillaume Duhamel, concepteur-rédacteur au groupe radio RTL avec lequel nous nous sommes entretenus :

> « Aujourd'hui tu as des animateurs qui se ressemblent tous, qui font tous les mêmes genres de *speak*, avec des contenus qui sont de plus en plus promotionnels. Je trouve qu'on a de moins en moins l'animateur de radio passionné de radio. Aujourd'hui les animateurs sont tous formés à la même école — le Studec — et sont formés à faire leurs *speaks* d'une certaine façon, et n'ont pas forcément une grande culture musicale. Des personnes comme Georges Lang, Francis Zégut ou Fred sur Skyrock, tu sens qu'ils ont une vraie culture. Quand ils parlent d'un artiste, d'un album ou d'un concert, il y a une vraie valeur-ajoutée. Et ça c'est l'ADN même de la radio, où il n'y avait pas toutes les contraintes publicitaires et promotionnelles d'aujourd'hui »[69].

La voix de radio concentre donc des effets que nous résumons sous le terme d'authenticité, qui suppose un savant dosage du discours entre la volonté de créer une intimité avec un auditeur tout en s'adressant à une communauté d'écoute. L'authenticité du média radio est palpable à toutes les échelles — chez le locuteur, chez l'auditeur, et auprès des publics. Que ces derniers considèrent la radio comme le média le plus crédible quant à la vérité de ses informations n'apparaît pas comme un hasard (à 56%, contre 52% pour la presse et 48% pour la télévision)[70]. L'authenticité du média radio est préservée par les gens qui la façonnent, des travailleurs du son aux directions des stations. Vladimir Bobrow, régisseur-monteur chez RTL, nous confiait en entretien que son travail de la voix répondait au positionnement de sa station, axé sur la proximité avec les publics :

> « Aujourd'hui RTL table en termes de couleur d'antenne sur quelque chose de très proche des gens. La proximité est le maître mot : ça reste toujours la radio de la famille, transgénérationnelle. (…) Aujourd'hui ce que j'adore dans le fait de travailler en radio, c'est qu'on se demande sans cesse "qu'est ce que la voix provoque chez les gens ?". Des voix très singulières comme celle de Vincent Perrot [animateur à la tête des petits matins de RTL et de l'émission musicale "Stop ou Encore" le week-end] fédèrent des communautés parce que ce sont des voix que les gens connaissent. Parfois dans la rue, tu vas marcher, entendre quelqu'un parler et te dire "mais attends je connais cette voix !". Ton cœur, ton

---

69 Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

70 Kantar France Insights, « Baromètre 2018 de la confiance des Français dans les Médias », 23/01/2018. Disponible à l'adresse suivante : https://fr.kantar.com/médias/digital/2018/barometre-2018-de-la-confiance-des-francais-dans-les-media/

corps, ta tête vont réagir de manière plus ou moins synchronisée ou pas selon que tu es content ou pas d'entendre cette voix. La voix est porteuse d'informations »[71].

C. LA VOIX, UNE OPPORTUNITÉ DE CRÉER DE NOUVELLES SITUATIONS D'INTERACTION ET D'IMMERSION AVEC LES PUBLICS DANS UN CONTEXTE DE VIRTUALITÉ NUMÉRIQUE.

1. La voix en contexte de virtualité : l'aboutissement des mécanismes d'immersion et d'interaction.

Le matériau sonore traverse l'histoire des technologies, et ce à plus d'un titre. Il est d'abord ce qu'elles ont désiré transporter. L'enregistrement et sa diffusion dans un phonographe ont marqué une première étape, suivie par le téléphone et la radio. Il est ensuite un matériau dont les propriétés ont été imitées. L'interaction et l'immersion permises par le son, et par extension la voix, sont en effet caractéristiques de l'économie numérique, qui se réalise dans un contexte de virtualité — le virtuel désignant ce « qui possède, contient toutes les conditions essentielles à son actualisation »[72]. Le son s'exprime dans un espace. Il crée une atmosphère dans laquelle l'auditeur est immergé, et provoquant en lui les effets que nous avons détaillés précédemment ; l'usage de la voix instaure une médiation de nos émotions et de nos intentions que l'on peut moduler dans sa dimension sonore, dans une situation d'interaction avec un auditeur ou un interlocuteur. La virtualité consiste en la création d'une atmosphère, d'un monde numérique dans un premier temps complément et désormais équivalent du monde physique ; elle propose également des modes d'interaction qu'elle figure en des « petites formes » visuelles comme les icônes, qui fonctionnent par analogie avec un objet physique[73], ou en incorporant une dimension sociale organisée en réseaux.

La voix et la virtualité partagent deux dimensions, la réduction de l'espace et du temps, que nous allons traiter tour à tour (a et b), afin de définir le statut de la voix en contexte numérique, dont

---

71 Entretien avec Vladimir Bobrow, régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 5, pp. 161-163.

72 *Trésor de la Langue Française Informatisé*, définition de « virtuel ».

73 Candel, E., Jeanne-Perrier, V., Souchier, E., « Petite Formes, Grands Desseins. D'une Grammaire des Énoncés Éditoriaux à la Standardisation des Écritures », in Davallon, J., (dir.), *L'Économie des Écritures sur le Web*, éd. Hermès-Lavoisier, 2012, pp.165-201.

les enceintes connectées sont une incarnation et qui constituent pour la radio un enjeu de redéfinition.

a) La voix et la virtualité, opérations de réduction de l'espace.

Les propriétés de la voix, qui provient du corps sensible, génèrent une immersion que le développement technologique a permis de déployer à l'échelle des sociétés. La voix est pour le corps une manière de s'approprier l'espace en extériorisant sa conscience, et en établissant un contact à la fois sensible (provenant de ses sonorités) et intelligible (lorsqu'elle est articulée dans un langage) avec les autres. Nous comprenons ainsi en quoi, outre le fait de pouvoir communiquer à une échelle géographique supérieure, le transport de la voix dans des technologies produit l'effacement de la distance. Pour Frances Dyson, qui dans son ouvrage *Sounding New Media* se lance dans l'ambitieuse entreprise de retracer l'histoire des technologies sous le prisme des propriétés du son, le téléphone a instauré la possibilité d'une « téléprésence » que la radio a prolongée en créant un « protocyberespace » dans lequel « les auditeurs pouvaient se voir enveloppés, absorbés, et parfois possédés » par le son à distance[74]. En radio, cette conquête de la distance trouve un symbole : le microphone, outil technique resserrant la distance réelle qui sépare un animateur de son auditeur. La distance entre l'animateur radio et son micro est en effet si minime que la proximité du studio se retrouve retranscrite à l'écoute du programme, de manière à la fois spirituelle et physique : *[« L'acte physique selon lequel la distance usuelle entre la source du son et le microphone est minime, devient une condition naturelle de l'art de la radiodiffusion comme une proximité spirituelle et atmosphérique entre l'animateur et l'auditeur »]* :

> « The physical fact that the normal distance between sound-source and microphone is inconsiderable, implies as a normal condition of the art of broadcasting a spiritual and atmospheric nearness of broadcaster and listener »[75].

La proximité qu'instaure le média radio est, comme nous l'avons vu, également travaillée par l'animateur dans son « phonostyle », pouvant lui conférer un effet d'intimité.

En outre, la radio réduit la distance en créant un effet de réel. La voix étant l'un des uniques matériaux utilisés pour transmettre l'information et l'émotion, elle construit sa réalité propre, qui

---

74 Traduit de l'Anglais. Dyson, F., *Sounding New Media. Immersion and Embodiment in the Arts and Culture*, éd. University of California Press, 2009, p.45.

75 Traduit de l'Anglais. Arnheim, R., *Radio*, éd. Da Capo Press, 1972, p.78, cité par Dyson, F., *Sounding New Media. Immersion and Embodiment in the Arts and Culture*, éd. University of California Press, 2009, p.53.

immerge l'auditeur dans un monde onirique, où il se crée ses images mentales. Analysant les récits radiophoniques de Pierre Bellemare, Anne-Marie Houdebine, chercheuse à l'université d'Angers, note sa capacité à construire un réel rien qu'avec sa voix : « PB signe de sa voix ce réel qu'il construit de sa voix , et ainsi il entraîne l'auditeur à le croire (« on s'y croit ») »[76]. Le travail de la voix en radio permet par conséquent l'immersion de son auditeur dans un monde sonore, faisant tomber les barrières de sa médiation.

La virtualité fonctionne de façon similaire, puisqu'elle consiste en une opération d'immersion des utilisateurs dans une autre réalité, non physique, grâce à la création d'un nouvel espace. A la manière de la radio, nous sommes immergés dans cet espace virtuel de manière immédiate, ou du moins par des modes d'accès et d'actualisation qui se veulent ergonomiques. Les icônes d'un bureau d'ordinateur, en présentant des analogies à des objets physiques (par exemple, la disquette symbolise l'enregistrement, la corbeille permet de supprimer des documents), facilitent la relation à la machine support de virtualité. Plus largement, les fonctionnalités développées sur le numérique nous permettent d'agir directement sur le contenu qui est affiché. Lev Manovitch, théoricien des nouveaux médias, considère par exemple que « les actions de « copier/coller » produisent des assemblages de mots, sons et images en syntonie avec les valeurs de l'esthétique post-moderne, (…) à l'heure actuelle, les techniques de composition permettent d'aller au-delà de la simple juxtaposition de données et de créer des « montages spatiaux » où des éléments hétérogènes se fondent selon les principes d'une nouvelle « esthétique de la continuité » »[77]. L'utilisateur est immergé dans un espace virtuel avec lequel il peut interagir dans une situation de proximité. Cette fonction d'interaction est présente en radio puisque l'auditeur réagit au contenu qu'il écoute en le projetant dans son imaginaire et ses émotions, mais pas directement (l'écoute demeure passive).

De nouvelles technologies visuelles intègrent l'immersion dans la narration même du contenu. Les exemples les plus édifiants sont le cinéma 3D et la Réalité Virtuelle, qui insèrent la personne dans le contenu même, lui conférant un pouvoir de préhension — que Isabelle Rieusset-Lemarié qualifie de « regard haptique », empruntant la terminologie de Deleuze[78]. La possibilité du toucher est avec ces technologies incorporée dans la vision.

---

76 Houdebine, A.-M., « Le récit radiophonique et son écoute : un procès de mythification », in Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984, p.56.

77 De Iulio, S., « Lev Manovitch, *The Language of New Media* », in *Questions de Communication*, 4, 2003, pp.473-475. Disponible à l'adresse suivante : https://journals.openedition.org/questionsdecommunication/5949

78 Rieusset-Lemarié, I., « Entre la « main de l'œil » et l' « œil digital », proximité et profondeur : la dimension haptique à l'horizon du Cinéma 3D et des acteurs virtuels », *Entrelacs*, 2013. Disponible à l'adresse suivante : https://journals.openedition.org/entrelacs/538

Le son, dont la voix constitue l'une des traductions et des composantes, partage avec la virtualité une capacité d'immersion dans un espace. Cette dernière est complétée d'une possibilité d'interaction moins prégnante dans le cas de la radio. Les enceintes connectées, en ce qu'elles proposent d'agir sur le contenu sur un mode impératif, où les ordres vocaux de l'utilisateur trouvent leur résolution dans la réponse de l'assistant intelligent, conjuguent ces qualités d'immersion et d'interaction dans un contexte à la fois virtuel et sonore. Ce contexte présente en outre la spécificité d'effectuer des allers-retours entre monde physique et monde virtuel, par la médiation d'un matériau corporel — la voix —, qu'une technologie s'approprie à son tour dans une logique de conversation.

b) L'instantanéité de la voix de radio et de la virtualité numérique.

La radio ne partagerait par le truchement de son matériau vocal qu'une qualité d'immersion, que le numérique — qui est une concrétisation du virtuel —, complète d'une interaction. Ce serait oublier la possibilité d'intervention des auditeurs sur le contenu radiophonique, qui peuvent prendre la parole en téléphonant au standard d'une station, et désormais réagir en direct à une émission sur les réseaux sociaux. La radio fonctionne par ailleurs historiquement en direct, et est en cela dotée de la même instantanéité que le numérique. Internet constituerait de ce fait une concrétisation virtuelle de qualités déjà à l'œuvre en radio depuis des décennies. Jérémy Valet, co-directeur de M6 Unlimited et auparavant directeur du pôle création des Opérations Publicitaires Spéciales du groupe radio RTL pendant dix ans, dessine un parallèle entre la radio et les réseaux sociaux, qui forment une osmose provenant du partage de certaines propriétés :

> « Les réseaux sociaux fonctionnent beaucoup mieux avec la radio que la diffusion *live* sur le site, et ce pour une simple et bonne raison : on dit que les réseaux sociaux c'est hyper moderne, et c'est vrai, mais il y a beaucoup de codes des réseaux sociaux qui sont compatibles avec la radio, qui existe depuis bien plus longtemps. Le fait de pouvoir être dans l'instantanéité de l'info. C'est-à-dire une info vient de tomber, et tout de suite on fait une brève, on fait une chronique, on en discute, on en débat. C'est quelque chose qui s'est toujours fait en radio, et c'est ce qui aujourd'hui existe avec les réseaux sociaux. Le principe de communauté, c'est-à-dire avoir une émission sur laquelle il y a une communauté qui écoute, qui se ressemble, qui partage des avis. Ça se passe sur les réseaux sociaux, et ça se passe aussi en radio avec des émissions de libre antenne sur lesquelles on posait une question »[79].

[79] Entretien avec Jérémy Valet, co-directeur de M6 Unlimited, mené par Clément Mellouet. Annexe 14, pp.194-204.

Nous pourrions par conséquent comparer les qualités de la radio à celles d'Internet, qui partagent leur instantanéité et leur capacité à fédérer des communautés dans l'espace. Le second constituerait l'apogée de la première, car il en élargit et en accélère les propriétés. Marshall McLuhan, en utilisant l'expression de « village global » engendrée par l'accélération technologique que portent les médias radio et télévision, illustre ce point :

> « L'accélération que nous connaissons aujourd'hui n'est pas une lente explosion du centre vers la périphérie, mais une implosion instantanée et une fusion de l'espace et des fonctions. Notre civilisation spécialisée et fragmentaire à structure centro-périphérique subit une réorganisation instantanée de tous ses fragments spécialisés en un tout organique. C'est le monde du village global »[80].

Cette citation résume en outre la manière dont le matériau vocal s'est déployé avec le progrès technique : une interaction entre voix intérieure et monde extérieur s'est vue conjuguée de la possibilité d'un dialogue, pour finalement être propagée et accélérée par la radio, puis par le numérique. Les communautés que fédèrent les émissions de radio fondent aujourd'hui des « communautés virtuelles »[81] sur les réseaux sociaux, et présentent la même capacité de réaction instantanée.

2. Un contexte porteur : entre « hyperesthésie » et saturation visuelle.

Sur le numérique, la radio s'est jusqu'à présent vue contrainte d'évoluer en réaction aux usages proposés par le numérique, majoritairement visuels. L'ajout d'images sur des séquences d'émissions afin de les partager sur les réseaux sociaux dénature sa fonction d'accompagnement — qui nous l'avons vu se déploie dans un dispositif acousmatique centré sur la voix —, en fragmentant le flux dans lequel elle s'exprime pour la rendre virale : « (…) ce ne sont pas des émissions en entier qui se partagent, mais des « morceaux » de radio. Surtout quand ceux-ci ont un fort pouvoir de « buzz »: une chronique de François Morel ou de Canteloup »[82]. Des initiatives graphiques pour figurer l'audio de manière visuelle ont été lancées, sans remédier tout à fait à ce problème. L'audiogramme, qui restitue les oscillations sonores d'un extrait audio, en est un exemple.

---

80 McLuhan, M., *Pour Comprendre les Médias*, éd. Seuil, 1968, p.117.

81 Rheingold, H., *Les Communautés Virtuelles*, éd. Addison Wesley France, 1996.

82 *Slate*, Kroh, L.,« Pourquoi est ce que la radio n'est pas virale ? », 28/02/2014. Disponible à l'adresse suivante : http://www.slate.fr/story/84009/radio-virale

Si les usages du numérique semblent contraindre la radio, nous avons vu qu'ils partagent avec elle des propriétés d'immersion et d'interaction qui se conjuguent en contexte de virtualité. Le numérique constituerait le sacre de l'image, disséminée sur les réseaux sociaux et les plateformes de *streaming* vidéo, que l'usage de l'audio vient compléter. Ses qualités d'immersion sont utilisées dans des narrations multimédias ; les podcasts, qui s'expriment sur le temps long, apaisent le rythme effréné de l'expression numérique ; les enceintes connectées utilisent la voix pour proposer de nouvelles interactions. Le numérique engendrerait en cela davantage un déclin du texte lu que de l'audio, et est en cela parfois décrit comme l'avènement d'une ère du « post-texte » : *[« Pensez aux innovations digitales les plus frappantes qui émergent aujourd'hui : les assistants vocaux qui ont été le succès des dernières vacances, le téléphone d'Apple qui scanne votre visage, l'Intelligence Artificielle qui permet de chercher des photos ou de traduire le langage parlé, et la réalité augmentée — qui insère de l'image digitale dans votre environnement immédiat. Ces avancées comprennent toutes des caméras, des microphones, votre voix, vos oreilles et vos yeux. Ensemble, elles nous envoient toutes le même message : bienvenue dans le futur post-texte »]* :

> « Consider the most compelling digital innovations now emerging: the talking assistants that were the hit of the holidays, Apple's face-reading phone, artificial intelligence to search photos or translate spoken language, and augmented reality — which inserts any digital image into a live view of your surroundings. These advances are all about cameras, microphones, your voice, your ears and your eyes. Together, they're all sending us the same message: welcome to the post-text future »[83].

Le contexte capitalistique dans lequel le numérique se développe, porté par les innovations de groupes valorisés à plusieurs de milliards de dollars représentés par les GAFAM (les géants de l'Internet : Google, Amazon, Facebook, Apple et Microsoft) a par ailleurs engendré une économie de l'attention mue par les sens de ses publics, une logique sensuelle. Un exemple : les communautés fédérées par l'ASMR (*Autonomous Sensory Meridian Response*), contenus audio et vidéos dont les sons procurent des frémissements au niveau de la tête et de la colonne vertébrale. Ce phénomène est né avec des contenus qui n'étaient pas destinés à procurer cet effet physique, catégorisés comme des « ASMR involontaires » (« *unintentional ASMR* ») sur des forums thématiques tels que Reddit, avant que des personnes ne produisent des contenus d'ASMR intentionnels, essentiellement sur YouTube. Des utilisateurs se sont ainsi appropriés des contenus à des fins sensorielles, décorrélées

[83] *The New York Times*, « Welcome to the Post-Text Future », 14/02/2018. Disponible à l'adresse suivante : https://www.nytimes.com/interactive/2018/02/09/technology/the-rise-of-a-visual-internet.html

de leur classification algorithmique sur les plateformes de vidéos. Rob Gallagher, un chercheur britannique, y voit une concrétisation de ce que David Howes nomme une « hyperesthésie » (« *hyperaesthesia* », qui en médecine définit l'exagération d'un sens) du capitalisme de fin de siècle[84]. Cette théorie est pensée comme un prolongement des travaux de Karl Marx sur le matérialisme, que Howes utilise pour désigner une société de consommation qui s'est déployée en mobilisant les sens des individus. Les réflexions sur les couleurs de l'emballage d'un produit pour attirer les consommateurs en sont une illustration. Les consommateurs traitent ainsi les marchandises de manière sensorielle et esthétique davantage que rationnelle. Rob Gallagher l'applique ainsi à l'ASMR :

> « La culture ASMR illustre la rapidité avec laquelle un sentiment diffus peut être converti en un certain type de vidéo capable de générer des revenus et souligne la dynamique affective et les fondements algorithmiques des logiques de consommation à une époque où, comme l'affirme David Howes, les marchés sont de plus en plus régis par la « logique sensuelle » »[85].

Si la culture ASMR s'est forgée de manière spontanée, à rebours des classifications algorithmiques, elle constitue désormais une économie insérée dans les fonctionnements des plateformes, au sein desquelles des amateurs produisent des contenus de « ASMR intentionnel ». La « logique sensorielle » qui anime cette communauté a ainsi été originée puis récupérée par les plateformes, qui en favorisent la viralité au sein de cette culture grâce à la personnalisation et au ciblage des algorithmes. Rob Gallagher poursuit cette réflexion en prenant l'exemple d'une vidéo à caractère humoristique devenue virale :

> « En 2006, l'une des premières vidéos sur YouTube à devenir « virale » montrait deux hommes en train de déposer des bonbons Mentos dans une bouteille de Coca-Cola pour créer un geyser gazéifié. La vidéo, qui déclencha sa propre éruption de vues, de partages, de commentaires et de vidéos imitatrices, est désormais un excellent exemple de la façon dont les plateformes d'édition en ligne permettent d'intégrer des actes de consommation et de production « hyper-esthésiques » dans une économie de l'attention qui (…) a engendré de nouveaux modèles commerciaux extrêmement lucratifs basés sur la publicité ciblée, la

---

84 Howes, D., « Hyperaesthesia, or The Sensual Logic of Late Capitalism », in *Empire of the Senses*, éd. David Howes, 2004.

85 Gallagher, R., « Générer l'euphorie en ligne : l'esthétique de la culture ASMR », in *Audimat*, vol.9, 2018, p100.

> sélection de contenu par l'algorithme, la fouille de données et l'analyse de sentiments »[86].

Les enceintes connectées s'insèrent à leur manière dans cette économie. Elle reposent sur le maniement d'un matériau sensible — la voix — qui procure une impression d'interactivité et d'immersion. Elles pourraient constituer en cela un double prolongement : celui des matériaux sensibles de la voix dans la technologie, et celui de l'interaction immédiate qu'elle permet dans des situations de consommation de contenus médiatiques ou de produits sur le numérique. Amazon, en créant son enceinte connectée Echo, a ainsi en tête de faciliter l'acte d'achat en ne mobilisant que la voix avec pour l'appuyer une immense base de données que lui procurent ses activités de e-commerce.

Par ailleurs, la voix pourrait plus simplement constituer un retour du temps long et de la profondeur dans une économie de l'attention numérique qui se caractérise par un effet de saturation visuelle. L'image est partout, et les consommateurs sur-sollicités apprécient le calme d'un contenu audio, propice à l'introspection (de par son effet d'intimité) et à l'imagination (grâce à son pouvoir évocateur). Pierre Laurent, chef de produit en charge de la création et de l'analyse sonore au pôle marketing du groupe radio RTL, l'exprime ainsi :

> « Aujourd'hui l'image est partout dans notre société. Que ce soit dans la rue, à la télévision, sur Internet avec les réseaux sociaux qui sont surpuissants, elle est au cœur de notre quotidien. C'est même la raison d'être de ces réseaux sociaux — Instagram, Snapchat, Pinterest... L'image est de plus en plus prégnante dans notre quotidien, qu'elle va jusqu'à saturer. Dans ce contexte le son a de vraies cartes à jouer. (…) Quand on n'écoute que le son, qu'on n'y appose pas d'image, [chacun] se fait sa propre interprétation de qu'il écoute. Sans l'image, tout devient plus grand, tout devient plus personnel »[87].

Nous avons, dans cette première partie de notre étude, d'abord exploré les propriétés de la voix. Son référent sensible trouve toute sa force lorsqu'il est incarné, et s'exprime dans les qualités musicales de la vive voix, qui entrent en tension avec un arbitraire du signe linguistique. Le transport et la propagation de la voix dans des ondes radio confèrent à ce média une authenticité, terme regroupant les effets d'intimité, de communauté, de réel et de projection dans l'imaginaire. Le numérique, qui constitue un nouveau champ de concurrence et d'expression pour la radio, partage avec la voix des propriétés d'immersion et d'interaction dans le temps et dans l'espace.

---

[86] ibid., pp.101-102.

[87] Entretien avec Pierre Laurent, Chef de produit chargé de la création et de l'analyse sonore au groupe radio RTL, mené par Clément Mellouet. Annexe 12, pp.187-191.

L'osmose entre la radio et les réseaux sociaux provient de cette ressemblance, qui se déploie dans un contexte de virtualité. Internet constitue en outre un contexte favorable à l'usage du matériau audio, car régi par une logique sensuelle propre à la société de consommation, ici projetée dans le monde virtuel. Si l'immersion que permet le matériau sonore, dont la voix est l'une des composantes, facilite son usage sur le numérique, c'est l'interaction immédiate qu'il instaure entre un utilisateur et son contenu qui cristallise l'intérêt d'une pluralité de médias pour les enceintes connectées. Ces dernières constituent par conséquent à la fois un prolongement et un aboutissement des qualités de la voix sur le numérique. Dans ce contexte, la voix se remédie. Elle change de forme, car elle n'est plus l'apanage de la radio, et parce qu'elle se trouve désincarnée par une Intelligence Artificielle, pouvant possiblement rompre les effets que procurent la vive voix. Nous allons à présent étudier l'altération — en tant que modification — de la voix sur les enceintes connectées, qui prend deux formes : le déplacement du matériau vocal d'un homme à une Intelligence Artificielle, et la construction d'une offre dans laquelle la radio perd son monopole du travail de la voix.

Nos considérations sur le partage de propriétés entre la voix et la virtualité numérique, s'il présage d'une harmonie et d'une efficacité des contenus qui les conjuguent, se heurtera en outre à la réalité d'un marché en construction. Deux limites s'érigeront ainsi dans notre démarche à teneur prospective. Premièrement, une description des acteurs en présence dans le marché des enceintes connectées fera montre d'une dépendance des entreprises médiatiques à l'implantation de cette technologie sur un territoire — l'offre médiatique sur Amazon Alexa en France, encore embryonnaire, permettra d'en attester. Deuxièmement, le déplacement de la voix d'un homme vers une machine demeure aujourd'hui imparfait. La figuration d'une personnalité artificielle, dont l'objectif est de réifier une présence corporelle, une proximité, est ainsi un idéal que les concepteurs des enceintes connectées tentent d'atteindre sans pour l'instant y parvenir.

II — L'économie de la voix numérique : champ concurrentiel élargi et altération artificielle.

Les enceintes connectées relèvent d'un phénomène singulier : la source de la voix se déplace d'un corps humain à une machine, alors que le propre de la voix est d'avoir ce *grain*, cette trace du vivant qu'elle incarne. Ce constat, quelque peu éthéré en l'état, le glissement d'un *grain* de la vive voix vers un *bruissement* de machine (pour reprendre les mots de Roland Barthes[88]) a des conséquences tout à fait concrètes pour l'ensemble des acteurs impliqués dans les enceintes connectées. Premièrement, ce nouveau marché provient d'une démarche d'innovation technologique extérieure aux médias qui y produisent des contenus. Les enceintes connectées sont exemplaires de la domination des GAFAM dans l'économie médiatique, éminence d'un processus marqué précédemment par les *Instant Articles* de Facebook ou le *Discovery* de Snapchat. Deuxièmement, cette désincarnation de la voix suppose un important investissement technologique de la part de ses concepteurs. La voix artificielle étant moins bien perçue par les audiences, la seule justification à sa création est le désir d'ubiquité de ses créateurs, et le seul critère à son adoption complète est sa vraisemblance — une adoption partielle étant permise par sa fonction d'assistant qui simplifie le quotidien. Notre fascination pour l'automaticité de l'objet doit idéalement faire place à un caractère naturel de l'interaction avec la machine, à une fluidité de conversation, pour s'ancrer dans nos vies quotidiennes. Avec les enceintes connectées, la radio évolue donc dans un univers concurrentiel élargi à un ensemble de marques-médias proposant des contenus audio, et contrainte par une voix artificielle qui encadre son expression. Nous allons à présent étudier les modalités du marché des enceintes connectées, contexte dans lequel la radio s'insère, puis leurs implications — l'organisation d'une concurrence élargie, et l'altération de la vive voix par une Intelligence Artificielle.

---

[88] Barthes, R., *Le Bruissement de la Langue*, éd. Seuil, 1984, p.99.

A. LE MARCHÉ DES ENCEINTES CONNECTÉES : L'INNOVATION MÉDIATIQUE SOUMISE À LA DOMINATION TECHNOLOGIQUE DES GAFAM.

1. État des lieux des acteurs en présence dans le marché des enceintes connectées.

Dans les quelques pages qui suivent, nous tenterons de décrire le marché des enceintes connectées, au sein duquel évolue la radio : son origine américaine, mise en perspective par l'émergence d'autres acteurs provenant d'Asie (a) et son degré de développement en France, territoire qui concentre notre intérêt (b).

a. Un marché provenant des Etats-Unis dont les rapports de force se réorganisent à l'échelle internationale.

Le marché des enceintes connectées est dominé en Occident par trois groupes nord-américains issus de l'économie numérique : Amazon, Google et dans une moindre mesure Apple. Les trois GAFAM se sont développés sur Internet, dont ils proposent un nouvel usage avec cette technologie. La force principale de chacun de ces groupes sur le numérique est aussi ce qui fait l'atout de leur enceinte connectée, au sein de laquelle une Intelligence Artificielle (Amazon Alexa, Google Assistant, et SIRI de Apple) s'associe à un objet (Amazon Echo, Google Home, et le HomePod). Les enceintes connectées Amazon Echo (qui composent une gamme de quatre modèles, du plus petit aux deux plus grands, munis d'un écran LCD : Echo Dot, Echo, Echo Spot et Echo Show) facilitent grandement l'achat en ligne car elles sont reliées aux services de e-commerce de leur marque-mère. Le Google Home (qui a aussi une version Google Home Mini, à l'instar de l'Amazon Echo Dot, et une version Max, proche de l'enceinte d'Apple) est le plus efficace pour les requêtes nécessitant le moteur de recherche de la firme éponyme, et est doté de l'Intelligence Artificielle aux plus grandes capacités de compréhension d'après Nicolas Dot, chef de produit marketing au groupe radio RTL en charge d'une veille sur les enceintes connectées[89]. Le HomePod d'Apple a une Intelligence Artificielle de moindre qualité, bien qu'elle ait été proposée aux consommateurs dès 2011 dans les iPhone 4S, et met en avant les qualités sonores de son produit ainsi que sa facilité de connexion aux autres appareils Apple.

---

[89] Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

La marque à la pomme est en retard dans le marché des enceintes connectées : elle n'a lancé son enceinte qu'en février 2018 aux États-Unis, contre respectivement novembre 2014 et 2016 pour Amazon et Google. L'entreprise de Jeff Bezos faisait ainsi office de précurseur lorsqu'elle a proposé Alexa sur le marché, devançant Google de deux ans. Cet ordre chronologique est retranscrit dans les parts de marchés de chacune de ces marques aux États-Unis, berceau des enceintes connectées : Amazon domine avec 61,9% de part de marché en mai 2018, contre 26,9% pour Google et 4,1% pour Apple[90]. La croissance d'Amazon est cependant ralentie par la montée en puissance de Google, qui est passé de 8,5% à 26,5% de part de marché entre janvier et mai 2018 aux États-Unis, quand le géant du e-commerce a perdu environ 8% de part de marché dans le même temps[91]. Cette baisse de régime donne lieu à la mise en place de stratégies de compensation, à défaut de pouvoir conserver la même hégémonie dans les ventes d'enceintes. Amazon a ainsi annoncé la mise en place d'une collaboration avec Microsoft, qui ne propose pas d'enceinte connectée mais une Intelligence Artificielle appelée Cortana sur ses tablettes et ordinateurs. Il suffira désormais de solliciter oralement l'une des Intelligences Artificielles — Amazon Alexa ou Microsoft Cortana — sur les appareils des deux marques, et ce de manière interchangeable. L'objectif est de profiter des atouts technologiques de chacune : « l'interaction profonde de Cortana avec Outlook pour les courriels et le calendrier ; les milliers de compétences et le contrôle domotique pour Alexa »[92].

En retard dans le marché des enceintes connectées, Apple joue la carte du contre-positionnement. Quand Amazon et Google mettent en avant la domotique (désignant l'usage de la technologie pour relier les fonctionnalités d'un foyer, allant des objets de communication numérique à la gestion de l'énergie) de leurs enceintes connectées, Apple met en avant la qualité acoustique de son appareil, censée justifier son prix très élevé — 350 euros contre en moyenne contre 100 euros pour Amazon Echo et 130 euros pour Google Home[93]. Un comparatif entre un film

---

90 *Voicebot*, Kinsella, B., « US Smart Speaker Market Share : Apple Debuts at 4.1%, Amazon Falls 10 Points and Google Rises », 03/06/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/06/03/u-s-smart-speaker-market-share-apple-debuts-at-4-1-amazon-falls-10-points-and-google-rises/

91 ibid.

92 *CNet France*, Zaffagni, M., « Amazon et Microsoft unissent leurs assistants Alexa et Cortana », 16/08/2018. Disponible à l'adresse suivante : https://www.cnetfrance.fr/news/amazon-et-microsoft-unissent-leurs-assistants-alexa-et-cortana-39872425.htm

93 *Le Monde*, Six, N., « Comparatif : HomePod, Echo, Home… Quelle enceinte connectée choisir ? », 23/06/2018. Disponible à l'adresse suivante : https://www.lemonde.fr/pixels/article/2018/06/23/comparatif-homepod-echo-home-quelle-enceinte-connectee-choisir_5320100_4408996.html

publicitaire de Google Home[94] et un d'Apple HomePod[95] est en cela édifiant. Quand Google Home est présenté dans une pluralité de situations domestiques (allumer les lumières du foyer ou apporter une précision dans une recette de cuisine), le HomePod n'apparaît qu'à deux reprises, laissant la place à un univers onirique où le foyer est cette fois-ci déconstruit (les murs se déplacent au gré de la danse de l'utilisatrice) ouvert par l'écoute d'une musique. Le film est d'ailleurs mis en scène par un réalisateur oscarisé, Spike Jonze, prolongeant son ancrage dans l'art. Ces captures d'écran soulignent l'écart entre les approches de Google et d'Apple :

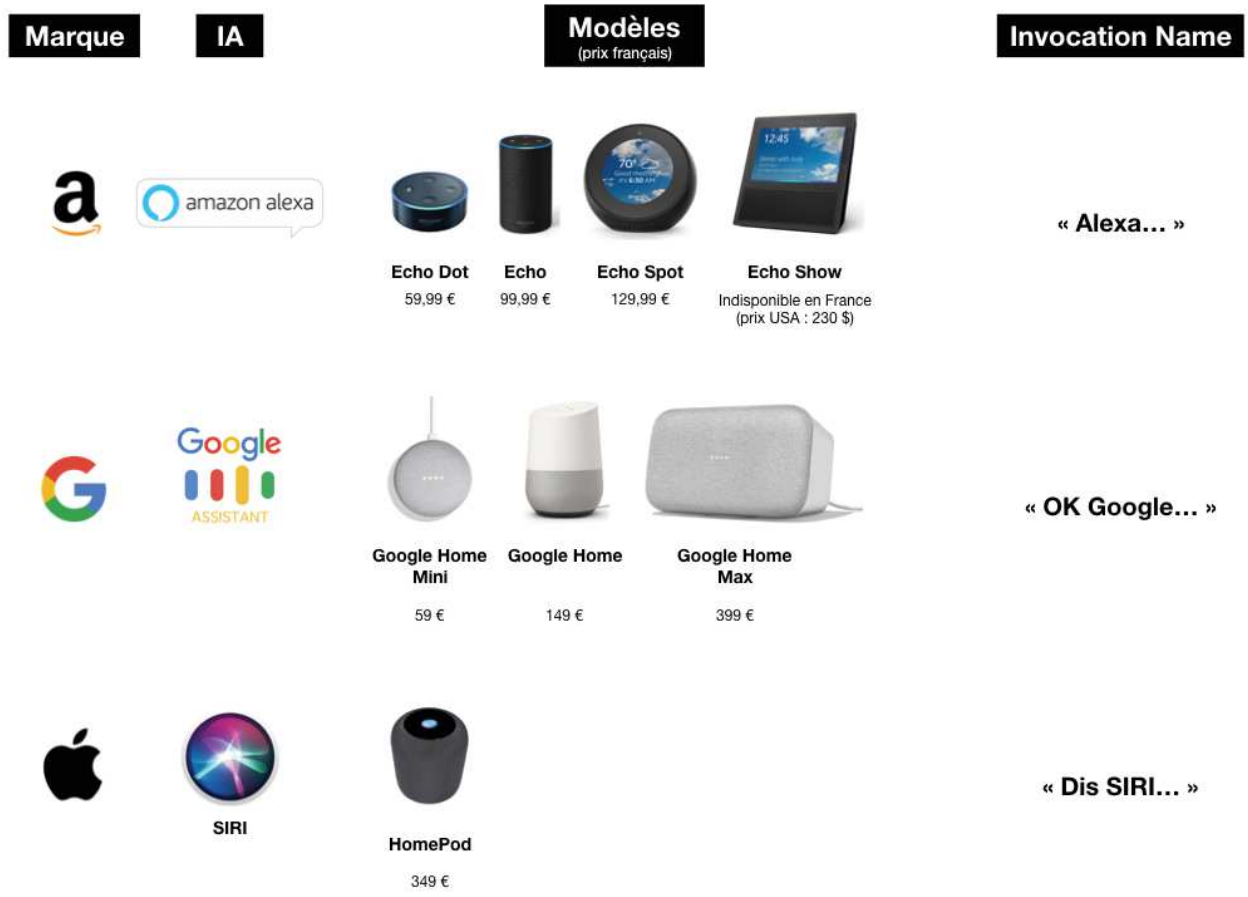




94 *Culture Pub*, « Google Home - États-Unis, 2017 ». Disponible à l'adresse suivante : « http://www.culturepub.fr/videos/google-google-home-google-home/

95 *Culture Pub*, « Apple : Welcome Home - États-Unis 2018 ». Disponible à l'adresse suivante : http://www.culturepub.fr/videos/apple-homepod-welcome-home/

Afin de récapituler les points que nous avons avons évoqués, le schéma ci-dessous[96] offre une illustration des modèles d'enceintes connectées proposés par Amazon, Google et Apple comprenant leurs prix sur le marché français ainsi que leurs « *invocation name* », expression désignant la phrase à laquelle ils s'activent :



Ces rapports de force entre Amazon, Google et Apple, initiés au sein du marché américain, ne se retranscrivent que partiellement à l'échelle internationale. Des entreprises asiatiques, puissantes sur leurs territoires mais peu implantées en Occident, et des marques aux publics moindres telles que l'enceinte connectée Djingo d'Orange en France, y représentent une partie non négligeable des ventes. Les deux marchés les plus puissants en termes de ventes après les États-Unis et leurs 4,1 millions d'unités vendues au premier quadrimestre 2018 sont la Chine (1,8 million) et la Corée du Sud (730 000) où dominent des marques nationales[97]. Dans le monde,

[96] Schéma réalisé par Clément Mellouet.

[97] *ZD Net*, Aubert, P., « Google Home passe devant Amazon Echo », 29/05/2018. Disponible à l'adresse suivante : https://www.zdnet.fr/blogs/watch-it/google-home-passe-devant-amazon-echo-39868815.htm

l'avance d'Amazon est contrebalancée par Google, déployé dans de nombreux pays, et ces acteurs asiatiques, portés par leurs territoires populeux. Ces acteurs emploient pour certains des stratégies agressives pour intégrer ce marché auparavant dominé par Amazon. Pour exemple, Google offrait sa Google Home Mini, lancée le 20 octobre 2017 en France, dans de nombreuses enseignes culturelles ou électroniques au bout d'un certain montant d'achat, ou lors d'occasions particulières[98]. Une autre étude rapporte des chiffres différents mais illustrant le même rapport de force : au premier trimestre 2018, Amazon aurait vendu 4 millions d'enceintes connectées dans le monde, Google 2,4 millions, Alibaba (entreprise de e-commerce chinoise) 700 000, Apple 600 000, Xiaomi (société d'électronique, elle aussi chinoise) 200 000 et les marques de taille moins importantes ont écoulé 1,3 millions d'enceintes connectées lorsqu'on cumule leurs ventes[99]. Cette dernière catégorie comprend en outre les *partners*, qui proposent leurs propres enceintes en utilisant une Intelligence Artificielle tierce, principalement Google Assistant et Alexa d'Amazon. Pour exemple, JBL, Panasonic, Sony et Lenovo embarquent le Google Assistant dans leurs enceintes connectées, et les marques Sonos et Ultimate Ears intègrent Amazon Alexa. A l'échelle internationale, Apple semble par conséquent moins puissant que le chinois Alibaba, et Google dominerait Amazon en représentant 36,2% des ventes contre 27,7%[100]. Le marché des enceintes connectées présente des disparités selon les territoires, où ce ne sont pas toujours les mêmes acteurs qui dominent. Les trois pays constituant les bassins les plus puissants en termes de parts de marché à l'échelle du monde en 2018 sont les États-Unis (45,6%), la Chine (20%) et la Corée du Sud (8,1%) — le reste du globe représentant 26,3% des ventes[101]. Ramenée à l'échelle mondiale, la commercialisation d'enceintes connectées fait ainsi l'objet d'une concurrence accrue entre des groupes provenant d'une diversité de pays, comme ce diagramme circulaire des ventes au premier trimestre 2018 le démontre[102] :

---

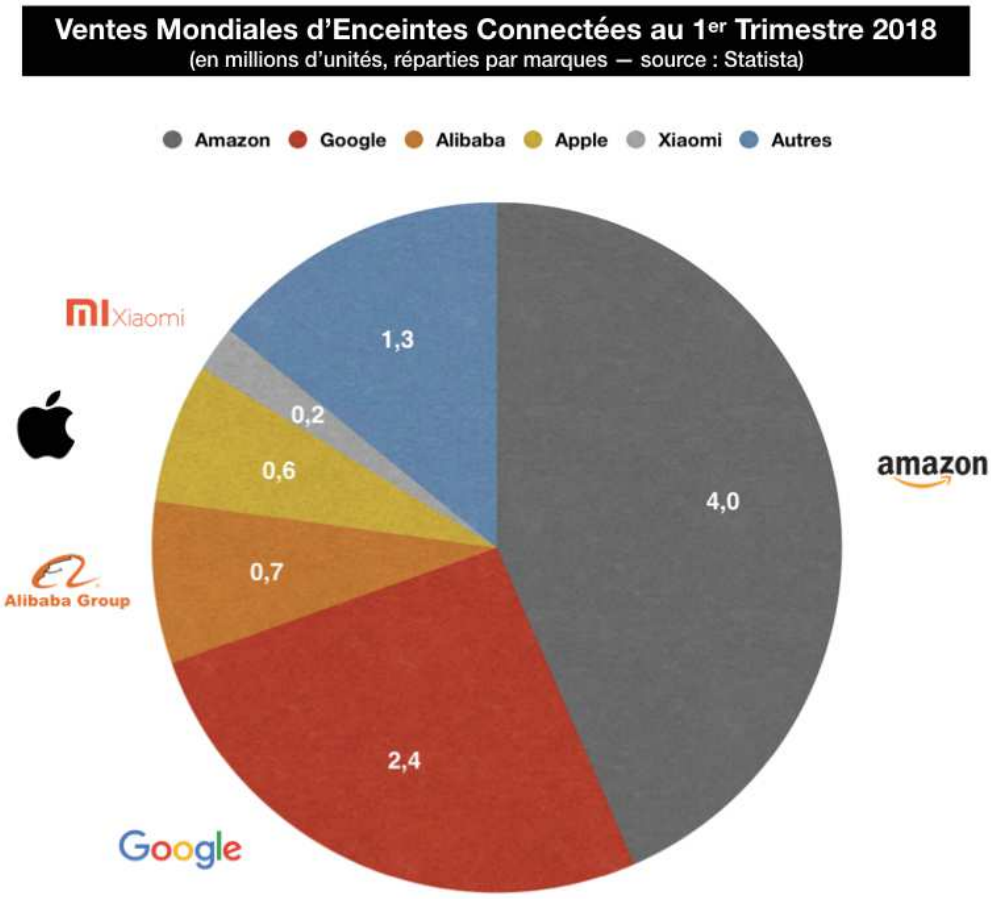
98 *iGeneration*, Moussie, S., « N'achetez pas le Google Home Mini… il est offert partout », 07/12/2017. Disponible à l'adresse suivante : https://www.igen.fr/accessoires/2017/12/nachetez-pas-le-google-home-mini-il-est-offert-partout-102247

99 *Statista*, Boittiaux, P., « Apple à la traîne sur le marché des enceintes connectées », 19/06/2018. Disponible à l'adresse suivante : https://fr.statista.com/infographie/12350/apple-a-la-traine-sur-le-marche-des-enceintes-connectees/

100 *ZD Net*, Aubert, P., « Google Home passe devant Amazon Echo », 29/05/2018. Disponible à l'adresse suivante : https://www.zdnet.fr/blogs/watch-it/google-home-passe-devant-amazon-echo-39868815.htm

101 *Voicebot*, Kinsella, B., « U.S. Falls to 46% of Smart Speaker Sales in Q1 2018 on the Rise of Alibaba, Xiaomi and Korean Device Makers », 23/05/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/05/23/u-s-falls-to-46-of-smart-speaker-sales-in-q1-2018-on-the-rise-of-alibaba-xiaomi-and-korean-device-makers/

102 Graphique réalisé par Clément Mellouet. Source : *Statista*, Boittiaux, P., « Apple à la traîne sur le marché des enceintes connectées », 19/06/2018. Disponible à l'adresse suivante : https://fr.statista.com/infographie/12350/apple-a-la-traine-sur-le-marche-des-enceintes-connectees/



Ce produit est en train de construire un secteur puissant, et les stratégies des acteurs qui le composent continuent de s'organiser. D'autres groupes d'envergure lancent leurs modèles, développant leurs propres Intelligences Artificielles (IA). Samsung a ainsi annoncé en août 2018 l'arrivée de Galaxy Home, utilisant son IA nommée Bixby déjà présente dans ses *smartphones* Galaxy S8 et S8+ depuis 2017[103]. Facebook a quant à lui freiné son projet d'enceinte connectée sous la pression du scandale Cambridge Analytica, symbole des menaces que le réseau social représente pour la vie privée de ses utilisateurs et que les enceintes connectées, capables d'enregistrer des conversations, auraient aggravé[104]. Si nous ne nous attarderons pas sur cet aspect qui n'est pas au cœur de notre analyse, remarquons que ces objets constituent une nouvelle intrusion

[103] *Les Numériques*, « Samsung Galaxy Home, l'enceinte connectée rivale du HomePod », 10/08/2018. Disponible à l'adresse suivante : https://www.lesnumeriques.com/assistant-domotique/samsung-galaxy-home-p45463/samsung-devoile-galaxy-home-enceinte-connectee-rivale-homepod-n77029.html

[104] *The Inquirer*, Page, C., « Facebook 'delays smart speaker launch' in wake of Cambridge Analytica scandal », 24/04/2018. Disponible à l'adresse suivante : https://www.theinquirer.net/inquirer/news/3030793/facebook-delays-smart-speaker-launch-in-wake-of-cambridge-analytica-scandal

des GAFAM dans l'intimité de leurs utilisateurs à des fins de ciblage publicitaire. Un couple de Portland, aux États-Unis, a par exemple eu la mésaventure de voir l'une de ses conversations enregistrée puis envoyée à l'un des contacts de son répertoire à son insu en mai 2018[105].

Les États-Unis, en tant que berceau du marché des enceintes connectées, sont aussi le pays où il est le plus développé. La pénétration de cette technologie dans les usages de la population américaine est en effet impressionnante. Quatre ans seulement après le lancement d'Amazon Echo, 18% des personnes âgées de 12 ans et plus possèdent une enceinte Google ou Amazon chez elles en 2018, soit 51 millions de personnes[106]. Parmi les possesseurs d'enceintes connectées, les premiers adopteurs (« *first adopters* » dans l'étude que nous citons) sont 28% à en avoir deux, et 27% à en compter trois ou plus dans leurs foyers ; dans la première couche du grand public (« *early mainstream* »), 30% possèdent deux enceintes connectées et 16% trois ou plus[107]. Au-delà de l'adoption des enceintes connectées dans leurs foyers, les personnes interrogées sont nombreuses à les utiliser de plus en plus dans le temps en comparaison avec le premier mois suivant leur achat — 48% pour les premiers adopteurs, et 54% pour la première couche du grand public[108]. Les enceintes connectées semblent par conséquent s'ancrer dans le quotidien de leurs utilisateurs, qui sont nombreux à songer à en acheter une autre pour leur maison ou leur appartement : 44% pour les premiers adopteurs (qui sont déjà nombreux à en posséder plusieurs) et 58% pour la première couche du grand public[109]. Comme le représente la publicité pour le Google Home que nous avons évoquée, les enceintes connectées font désormais partie de la vie des personnes qui les ont adoptées. Pour preuve, 25% des parents possédant une enceinte connectée déclarent qu'elles aident leurs enfants à faire leurs devoirs[110]. Cette avalanche de chiffres ne signifie finalement qu'une chose : les enceintes connectées sont un phénomène, allant jusqu'à créer une bulle spéculative dans laquelle leur sont prédits des sommets. Une étude de Juniper Research, cabinet d'études marketing britannique spécialisé dans le numérique, prévoit ainsi que l'on comptera 275 millions d'appareils munis d'assistants vocaux dans des maisons connectées d'ici 2023, soit une croissance de 1000%

---

105 *Oregon Live*, Acker, L., « Portland woman says Amazon Alexa recorded personal conversation, sent it to contact », 25/05/2018. Disponible à l'adresse suivante : https://www.oregonlive.com/trending/2018/05/portland_woman_says_amazon_ale.html

106 Edison Research, « The Infinite Dial 2018 », 08/03/2018, p.8. Disponible à l'adresse suivante : http://www.edisonresearch.com/infinite-dial-2018/

107 Edison Research & National Public Radio, « The Smart Audio Report - Spring 2018 », p.12. Disponible à l'adresse suivante : https://www.nationalpublicmedia.com/smart-audio-report/latest-report/

108 ibid., p.11.

109 ibid., p.13.

110 ibid., p.37.

par rapport à 2018[111]. Les enceintes connectées deviennent un enjeu pour les marques commerciales, qui peuvent y vendre leurs produits et leurs services par l'unique entremise de la voix et soutenues par un ciblage algorithmique, comme nous l'explique Nicolas Dot :

> « Les enceintes connectées, d'un point de vue commercial et publicitaire, permettent de récupérer un maximum de données, surtout vocales afin de pouvoir pousser des promotions et, au lieu de faire un achat par le *search*, faire un achat par la voix. Cela permettrait de faciliter encore l'acte d'achat en le raccourcissant et en le ciblant davantage, en fonction des habitudes d'écoute et des préférences d'utilisation de chacun »[112].

Si nous détaillerons les usages des enceintes connectées dans un second temps, nous pouvons d'ores et déjà noter que leur marché, de par sa croissance économique et l'interaction directe que son produit permet grâce à la voix, suscite l'intérêt des marques, qui comme nous le verrons y déploient leur présence. Le commerce vocal devrait par exemple se développer très rapidement, et générer 40 milliards de dollars à l'échelle du monde d'ici 2022, contre 2 milliards de dollars aujourd'hui[113]. Dans le cas des médias, ce déploiement s'effectue dans une logique d'ubiquité — de média global qui touche des publics sur une multitude de supports. Un indice de leur intérêt pour les enceintes connectées : l'étude que nous avons majoritairement citée pour décrire le marché américain, l'édition printemps 2018 du « Smart Audio Report », est le fruit de la collaboration de Edison Research (un cabinet d'études marketing) et de la radio publique américaine (*National Public Radio*, abrégé en NPR), qui en publient une version à chaque mi-saison depuis un an.

b. Un marché émergent en France.

Le marché français des enceintes connectées se développe dans un contexte européen, où les marques asiatiques que nous avons citées — notamment les chinois Alibaba et Xiaomi — ne sont pas implantées. Google et Amazon se partagent la majorité des ventes. Au Royaume-Uni, en

---

111 *Business Wire*, « Juniper Research: Voice Assistants Used in Smart Homes to Grow 1,000%, Reaching 275 Million by 2023, as Alexa Leads the Way », 25/06/2018. Disponible à l'adresse suivante : https://www.businesswire.com/news/home/20180625005474/en/Juniper-Research-Voice-Assistants-Smart-Homes-Grow

112 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

113 *PR Newswire*, « Voice Shopping Set to Jump to $40 Billion By 2022, Rising From $2 Billion Today », 28/02/2018. Disponible à l'adresse suivante : https://www.prnewswire.com/news-releases/voice-shopping-set-to-jump-to-40-billion-by-2022-rising-from-2-billion-today-300605596.html

Allemagne et en France, le premier rassemblait 33% de part de marché et le second 67% en 2017, pour un total de 6,5 millions d'unités vendues. Le Royaume-Uni apparaît comme le pays le plus réceptif à cette technologie, avec 3,94 millions d'enceintes connectées vendues contre 500 000 en France au sein de ce total[114]. Si l'arrivée récente de cette technologie dans l'hexagone explique en partie sa faible adoption, nous observons également un manque de compréhension quant à son utilité dans le foyer. Les résultats d'une étude menée par MindShare, agence média de GroupM, sur un échantillon de 825 hommes et femmes âgés de 16 à 70 ans, révèle la perception d'un manque de pédagogie auprès des publics : « 57% [des sondés] se disent intéressés par le sujet, mais 62% déclarent ne pas être assez informés et pensent à 38% que plus d'informations sur les assistants vocaux seraient un levier pour accroître leur utilisation »[115]. Au regard des États-Unis, les enceintes connectées représentent en France un marché embryonnaire mais, plus étonnement, un marché moins important que chez ses homologues européens. Seuls 6% des Français posséderaient une enceinte connectée d'après un sondage mené sur 1 458 personnes en mai 2018[116]. Une comparaison du nombre de *skills* — équivalent des *apps*, ou applications, sur Google Home — sur Amazon Alexa entre plusieurs pays illustre ce constat (déjà amorcé avec l'écart du nombre de ventes) : 25 784 aux États-Unis, 8 982 au Royaume-Uni, et 3 118 en Allemagne en janvier 2018[117], pour environ 200 *skills* en langue française[118].

Notre analyse du marché des enceintes connectées, que nous avons pour terminer resserrée à la France, atteste de la domination des GAFAM dans l'offre médiatique et commerciale présente sur ce support. A l'origine de la technologie, le degré de leur implantation est une condition au déploiement de l'offre de contenus, comme semble le prouver le faible nombre de *skills* sur la version française d'Amazon Alexa. La richesse de l'offre sur les enceintes connectées ne répond cependant pas uniquement à une notion de quantité. Les services propres des marques qui proposent la technologie sont présents par défaut sur les appareils, et ceux d'acteurs extérieurs sont mis en

---

114 *Voicebot*, Kinsella, B., « Amazon Echo and Google Home European Smart Speaker Sales Approach 6.5 Million Units in 2017 », 28/02/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/02/28/amazon-echo-google-home-european-smart-speaker-sales-approach-6-5-million-units-2017/

115 *Offre Média*, « Le potentiel du marché des assistants vocaux étudié par Mindshare », 27/04/2018. Disponible à l'adresse suivante : https://www.offremedia.com/le-potentiel-du-marche-des-assistants-vocaux-etudie-par-mindshare

116 *Stratégies*, « 6% des Français ont une enceinte connectée », 23/05/2018. Disponible à l'adresse suivante : http://www.strategies.fr/actualites/medias/4013057W/6-des-francais-ont-une-enceinte-connectee.html

117 *Voicebot*, Kinsella, B., « Amazon Closes Year With 266% Alexa Skill Growth in U.S. », 08/01/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/01/08/amazon-closes-year-with-266-alexa-skill-growth-u-s/

118 *Journal du Net*, « Echo / Alexa : l'enceinte d'Amazon à la loupe [prix, apps, config…] », 23/08/2018. Disponible à l'adresse suivante : https://www.journaldunet.com/ebusiness/internet-mobile/1194132-echo-alexa-l-enceinte-d-amazon-a-la-loupe-prix-apps-config/

avant par le biais de partenariats. En toute logique, Google Home ne fera pas une recherche web sur Bing, le moteur de recherche de Microsoft, ou sur celui du Français Qwant. Par extension, les données récupérées par les enceintes connectées sur leurs utilisateurs viendront nourrir leurs propres services de ciblage publicitaire. Des sociétés indépendantes — telles que les Français Zac (dédié à la domotique) ou XBrain, assistant vocal pour la voiture[119] — tentent pour cette raison de contrecarrer la tendance à l'hégémonie des GAFAM au sein du marché des enceintes connectées en proposant leurs propres produits.

L'implantation des enceintes connectées sur un territoire n'est néanmoins pas suffisante pour analyser les modalités de construction d'une offre médiatique. Il nous faut croiser cette donnée avec les usages des enceintes connectées, qui rassemblent une multiplicité de fonctions, et créent une situation de concurrence horizontale entre des marques médias, des entreprises commerciales, des fonctionnalités domotiques, et des plateformes numériques (par exemple de *streaming* musical, concurrent direct des radios spécialisées dans ce domaine).

2. Analyse des usages des enceintes connectées : des nouvelles situations d'écoutes pour la radio.

Son aspect de transistor, augmenté de la possibilité d'une interaction directe entre un utilisateur, des contenus médiatiques et des fonctionnalités, fait de notre objet bien plus qu'une enceinte. Elle est un assistant vocal, présent à chaque étape du quotidien. Si elles constituent un champ de création et de concurrence pour les contenus médiatiques, les enceintes connectées ne se destinent ainsi pas en premier lieu à leur consommation. L'étude « Smart Audio Report » menée par NPR (*National Public Radio*, radio publique américaine) et Edison Research (cabinet d'études marketing), dont nous avons déjà cité l'édition printemps 2018 plus tôt, apporte des données concrètes sur les usages des enceintes connectées. Ne réfléchissant plus en termes de vente de l'objet mais en termes d'usages, il nous faut préciser que cette étude a été menée en ligne auprès de 909 personnes âgées de 18 ans et plus, et qu'elle ne concerne que les États-Unis.

Au regard des chiffres présentés dans cette étude, nous constatons tout d'abord que l'écoute de contenus audio — rassemblant la musique, la radio AM et FM, les podcasts, les contenus dédiés aux enceintes et les livres audio — augmente avec l'achat d'une enceinte connectée : 63% des premiers adopteurs (« *first adopters* ») et 70% de la première couche du grand public (« *early*

---

119 *L'ADN*, Robillart, O., « On n'aura bientôt plus besoin d'écrire... La révolution de la voix est en marche », 02/02/2018. Disponible à l'adresse suivante : http://www.ladn.eu/tech-a-suivre/ia-et-chatbot/reconnaissance-vocale-les-start-up-en-france/

mainstream ») déclarent écouter plus de contenus audio depuis leur achat[120], rejoignant le fait que l'enceinte est de plus en plus utilisée avec le temps, comme nous l'évoquions précédemment. Au sein de ces contenus audio, la radio linéaire (le flux en direct ici diffusé par Internet) et délinéarisée (les podcasts de rattrapage ou natifs, c'est-à-dire dédiés à l'écoute digitale) n'est cependant pas le réflexe premier des utilisateurs. Les usages principaux des enceintes connectées (ici quantifiés sur une semaine) pourraient en effet être qualifiés d'utilitaires ou de fonctionnels, confirmant leur statut d'assistants du quotidien : la musique est l'usage le plus sollicité (90% des premiers adopteurs et 89% de la première couche du grand public ont répondu « oui » au fait qu'ils en consommaient sur ce support), devant la météo (81% et 85%), les questions d'ordre général (83% et 85%), la fonction minuteur (66% et 69%), l'heure (61% et 65%), écouter l'actualité, qui peut provenir de nombreux autres médias que la radio ayant développé une application dédiée (51% et 58%), et demander à l'assistant de raconter une blague (50% et 48%). Dans cette liste des usages, qui pourrait encore continuer, la radio en direct n'est déclarée être écoutée que par 41% des premiers adopteurs et 47% de la première couche du grand public. En l'espace d'une semaine, des podcasts ont estimé avoir été écoutés par 20% des premiers adopteurs et 28% de la première couche du grand public[121]. L'étude conclut ainsi que les premiers adopteurs « ont un usage plus avancé des enceintes connectées, concentré sur la sécurité de la maison et autres outils domestiques », et que la première couche du grand public les sollicite « pour un champ plus large d'activités quotidiennes — commander de la nourriture, passer des appels, obtenir l'information trafic, rechercher des produits et faire des achats »[122]. La consommation de contenus médiatiques n'est ainsi pas un réflexe, constat que Sacha Ephraim, régisseur-monteur chez RTL possédant une enceinte connectée, a corroboré durant notre entretien :

> « *[Qu'est ce que vous écoutez sur votre enceinte connectée ?]* De la musique. Et comme je n'ai pas de radio qui regroupe tout ce que j'aime en musique, je suis obligé de me faire ma propre petite radio personnelle. Des petites *playlists*... C'est bête, mais si vous ne me l'aviez pas souligné, je n'aurais pas mis la radio sur mon enceinte connectée. Ce n'est pas un réflexe automatique »[123].

---

120 Edison Research & National Public Radio, « The Smart Audio Report - Spring 2018 », p.33. Disponible à l'adresse suivante : https://www.nationalpublicmedia.com/smart-audio-report/latest-report/

121 ibid., pp.25-27.

122 Traduit de l'Anglais. Ibid., p.29.

123 Entretien avec Sacha Ephraim, Régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 6, pp. 164-167.

Un type de programme audio se détache pourtant du lot, et concentre les usages : l'actualité, qui rappelons-le est consommée par 51% des premiers adopteurs et 58% de la première couche du grand public possédant une enceinte connectée. Les premiers sont 24% à déclarer en écouter une à deux heures par semaine, et les seconds 33%[124]. Sans que l'étude ne précise les sources exactes de l'information que les personnes interrogées déclarent écouter, nous observons que le traitement de l'actualité est un objet médiatique et, sur ce support audio, éminemment lié aux usages de la radio. La place de la consommation d'actualité dans une typologie des usages des enceintes connectées au cours d'une journée proposée dans l'étude que nous citons n'est pas sans rappeler les audiences de la radio, qui se concentrent le matin — notamment durant le *prime* de 6h00 à 9h00 — pour décliner le reste de la journée (avec un léger rebond à midi et lors du *drive*, qui correspond à la fin de journée, lorsque les auditeurs sont à l'écoute de la radio sur le trajet à leur domicile) : l'information sur le trafic routier est la fonction la plus sollicitée par les utilisateurs entre 5h00 et 9h00 du matin, suivie par la météo et l'actualité[125]. La radio pourrait par conséquent avoir une carte à jouer sur les enceintes connectées, mais n'a pas développé un réflexe vers son média. Cela nous apparaît paradoxal car la voix, dont les propriétés d'immersion et d'interaction imprègnent les enceintes connectées, sont intimement attachées à l'identité du média radio. Sacha Ephraim rejoint ce constat :

> « Les radios ont sûrement encore du travail à faire pour affirmer leur présence sur ce support. Il y aurait quelque chose à faire d'intéressant pour suggérer aux gens que cet objet permet d'écouter des émissions de radio avec une simplicité déconcertante »[126].

Nicolas Dot, chef de produit au marketing radio du groupe RTL, abonde lui aussi en ce sens en précisant que l'ouverture que les enceintes connectées constituent pour le média radio pourrait vite se refermer face à la concurrence des autres médias :

> « C'est pour ça que la radio doit saisir cette opportunité, notamment en termes de podcasts pour exploiter ce marché qui est en pleine effervescence et associer ce marché à son média. Le danger de plus en plus pour la radio c'est qu'il y a beaucoup d'acteurs qui se penchent sur le podcast, et donc sur les enceintes

---

124 Edison Research & National Public Radio, « The Smart Audio Report - Spring 2018 », p.33. Disponible à l'adresse suivante : https://www.nationalpublicmedia.com/smart-audio-report/latest-report/, p.21.

125 ibid., p.32.

126 Entretien avec Sacha Ephraim, Régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 6, pp. 164-167.

> connectées. Ce sont des titres de presse, des groupes de télé... Surtout de presse. Ce qui pourrait créer une nouvelle concurrence pour la radio de ce point de vue des podcasts et des contenus dédiés aux enceintes. On l'observe déjà pour les podcasts, avec l'arrivée de podcasts natifs qui ont chamboulé l'ordre établi »[127].

Au-delà de cet avantage concurrentiel momentané, la radio pourrait avec les enceintes connectées modifier sa structure d'audience, qui en 2015 allait de 64 ans pour France Musique à 29 ans pour Skyrock en France sur la période septembre-décembre d'après *BFM Business*[128]. La réflexion ne peut cependant être si simple. Une radio, en développant des contenus pour des supports tels que les enceintes connectées, doit trouver le juste dosage entre son offre délinéarisée et celle de son flux en direct, qui demeure son *médium* de prédilection — un équilibre entre la cohérence de son offre, constitutive de sa marque-média et de son audience, et l'évolution de ses modes d'accès qui passe, pour créer un réflexe, par la création de contenus dédiés aux usages qu'ils proposent. Ce mouvement s'inscrirait dans une perspective d'anticipation de l'évolution des usages de génération en génération, qui se déplacent vers le numérique.

Les enceintes connectées constituent par conséquent une réminiscence du transistor, avec lequel la radio est historiquement écoutée, qu'elles complètent d'une possibilité d'interaction directe avec les utilisateurs. Sur ce support, les radios ont un avantage concurrentiel provenant de leur ancrage historique dans le maniement de la voix, matériau dont elles travaillent le référent sensible. Elles subissent parallèlement une double pression, qui contraint leur déploiement sur les enceintes connectées, dont allons étudier chaque aspect tour à la tour : la concurrence horizontale et directe de médias qui proposent des contenus audio dans une logique de média global qui multiplie les supports d'expression (partie B) ; puis la dépendance aux avancées technologiques des GAFAM, dont les « Intelligences » utilisent des voix artificielles, et représentent à la fois une contrainte et un avantage pour la radio, qui peut potentiellement faire ressortir son authenticité au sein de cet encadrement (partie C).

---

127 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

128 *BFM Business*, Henni, J., « Quelle est la radio dont les auditeurs sont les plus vieux? », 14/02/2016. Disponible à l'adresse suivante : https://bfmbusiness.bfmtv.com/entreprise/quelle-est-la-radio-dont-les-auditeurs-sont-les-plus-vieux-950088.html

B. ANALYSE DE L'OFFRE DES MÉDIAS FRANÇAIS PRÉSENTS SUR L'ENCEINTE CONNECTÉE AMAZON ECHO.

1. Mise en contexte : une stratégie de déploiement régie par la notion de « média global ».

Notre analyse du marché des enceintes connectées nous a permis de constater que l'implantation d'une technologie sur un territoire a une influence directe sur le déploiement d'une offre médiatique. Cela semble logique, puisque les médias se destinent à parler à un public suffisamment nombreux et qu'ils peuvent monétiser. L'absence d'une mesure d'audience instaurée et légitime complique la vente d'espace publicitaire à des annonceurs au sein des contenus médiatiques dédiés aux enceintes connectées. L'étude Global Radio de Médiamétrie, organisme officiel en charge des mesures d'audiences des médias en France, ne comprend pas les enceintes connectée dans sa mesure des écoutes digitales de la radio (à la fois le direct et le *replay*). Les catégories prises en compte et les publics qu'elles rassemblent sur la période janvier-mars 2018 sont les suivants : sur l'ensemble de l'écoute quotidienne du média, 12,2% (soit 6,6 millions de paires d'oreilles) est numérique, total au sein duquel le mobile arrive en tête (3,5 millions), suivie par l'ordinateur (1,7 million), la télévision (1,1 million), la tablette (505 000) et le baladeur (118 000)[129]. Parmi les méthodes pour insérer des annonceurs dans des contenus de médias sur les enceintes connectées — le *sponsoring* d'un programme, les écrans publicitaires classiques, et les contenus de marque —, celle de l' « *ad-switching* », sorte de compromis entre les écrans publicitaires et la personnalisation permise par les algorithmes semble être la voie privilégiée. Cette technique, déjà utilisée en TV numérique, consiste à « substituer un spot publicitaire diffusé dans le flux en direct d'une chaîne par un autre spot sur un support numérique, ordinateur, smartphone ou tablette »[130] pour en augmenter la couverture sur une cible commerciale. Elle ne peut trouver son efficience et sa vente à des annonceurs qu'en se fondant sur une mesure d'audience instituée, pour l'instant absente des contenus dédiés aux enceintes connectées. Nicolas Dot résume ainsi cette problématique :

---

129 *Offre Média*, « Les supports multimédias atteignent 12,2% du volume total d'écoute de la radio selon Médiamétrie », 18/05/2018. Disponible à l'adresse suivante : https://www.offremedia.com/les-supports-multimedias-atteignent-122-du-volume-total-decoute-de-la-radio-selon-mediametrie

130 *Stratégies*, Fraioli, B., « L'ad-switching à l'ère du médiaplanning », 20/04/2017. Disponible à l'adresse suivante : http://www.strategies.fr/actualites/medias/1062735W/l-ad-switching-a-l-ere-du-mediaplanning.html

> « Comment on monétise l'écoute via ces enceintes connectées ? Est ce qu'on choisit d'avoir un *ad switching* comme en TV sur le digital — avoir des publicités ciblées sur Internet, différentes de la publicité qui passe en télé — sur les enceintes ? C'est *a priori* ce vers quoi on se dirige, en tout cas pour des programmes comme les podcasts. Après, combien ça va représenter, à quel point va-t-on avoir accès la donnée ? En effet, on observe une cannibalisation par les GAFA, qui ont tendance à retenir les informations. Sans possibilité de mesurer de façon précise les audiences, un peu comme sur iTunes, la monétisation sera d'autant plus complexe. Il faut aussi que cet objet rentre dans les us et coutumes pour que la monétisation prenne de l'ampleur »[131].

La présence d'un nombre conséquent de médias français sur les enceintes connectées ne relève dès lors pas d'un enjeu de monétisation ou d'enrichissement dans cet état embryonnaire du marché. Il s'agit davantage d'affirmer leur présence dès les balbutiements de cette technologie pour grandir avec elle, dans l'espoir d'un essor semblable à celui du marché américain. Cette volonté se conjugue d'un enjeu d'image, très vif dans le cas de la radio, de média ancré dans la modernité numérique qui anticipe les évolutions des usages. Lorsqu'un média s'introduit sur Google Home, Amazon Echo ou le Apple HomePod, il construit son identité auprès des publics (une communication *Business to Customer*) et des annonceurs qui assurent sa rentabilité (*Business to Business*). Le média ne tire aucun gain financier de son association à l'enceinte connectée ; il dépense de l'argent pour créer des contenus spécifiques de la même manière que le développement d'une application mobile suppose d'y allouer des ressources. Pierre Chausse, directeur adjoint des rédactions du *Parisien*, déclarait ainsi qu' « aujourd'hui, on ne se rémunère pas (…). Il faut à l'inverse dépenser de l'argent (« pas beaucoup ») pour être sur Amazon Echo »[132]. Le rajeunissement de l'audience du média radio passe, avant sa conquête de publics plus jeunes, par un travail de l'image des marques qui le composent :

> « C'est une course à l'innovation, une volonté d'être les premiers à adopter un service le plus innovant possible. En dehors de cet aspect innovant, porteur en termes d'image pour les médias, c'est aussi un enjeu de rajeunissement et de diversification de l'audience qui est particulièrement important dans un écosystème médiatique qui n'a de cesse d'être chamboulé et perturbé par de nouveaux acteurs, toujours plus compétitifs. Pour mettre l'accent sur la radio, elle

---

131 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

132 *Numerama*, Turcan, M., « Comment Amazon Echo s'appuie sur les médias pour conquérir les Français », 06/06/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/380039-comment-amazon-echo-sappuie-sur-les-medias-pour-conquerir-les-francais.html

> joue sur cet effet d'annonce. Ce n'est pas pour rien que chaque média proclame être le premier présent sur Amazon Echo, sur Google Home »[133].

Les médias classiques — adjectif rassemblant de manière arbitraire la radio, la télévision et la presse, qui existaient avant et se remédient sur le numérique — s'insèrent dans une logique de média global, qui mélange image ancrée dans l'innovation et capacité à s'exprimer sur une multitude de supports. Cette stratégie suppose d'être capable de construire une ligne éditoriale assujettie à une double tension, entre la capacité à la conserver sur des supports distincts (le renvoi à une marque ancrée sur un support d'origine, en lien avec ses savoir-faire sources de sa valeur ajoutée) et de créer des contenus modernes en ce que leur expression apparaît comme naturelle sur le nouveau support. Mathieu Gallet, ancien président du groupe public Radio France, définit cet enjeu dans une note publiée sur le site du *think-tank* Terra Nova le 5 novembre 2017 :

> « Les médias classiques doivent évoluer pour agréger l'ensemble des usages et permettre une expérience plurielle à des utilisateurs qui accèdent à leurs contenus par de multiples canaux : le média global, c'est la mise à disposition de tous les programmes ou émissions sous toutes les formes possibles, en direct comme en différé, programmés ou à la demande, à l'écrit, en photo, en son, en vidéo, sur tous les réseaux et en particulier les réseaux numériques. Ainsi, le contenu fourni (information, fiction, documentaire, humour, réflexion...) est retravaillé pour être proposé sous toutes ses facettes médiatiques : enrichi, réfléchi, renforcé, et plus attrayant pour toutes les catégories de public »[134].

La radio d'information en continu franceinfo, citée plus loin dans la note, illustre cette volonté de déploiement et de signification du média d'origine. La radio, fondée en 1987, a ajouté à ses canaux d'expression un site web, une application et une chaîne de télévision. Le partage de programmes entre sa chaîne TV et sa radio — notamment les *flashes info*, ou la diffusion simultanée d'une partie de la matinale sur les deux supports — et leur reprise sur Internet fait circuler une image de marque ancrée dans la radio dans divers canaux d'expression. Cette stratégie de média global répond à un enjeu de diversification de l'audience et d'image, les deux ayant une influence réciproque. Pour

---

133 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

134 *Terra Nova*, Gallet, M., « Transformation numérique, synergies agiles, responsabilité démocratique : trois enjeux pour les médias publics », 05/11/2017. Disponible à l'adresse suivante : http://tnova.fr/notes/transformation-numerique-synergies-agiles-responsabilite-democratique-trois-enjeux-pour-les-medias-publics

exemple, franceinfo a ancré son image dans la modernité numérique en s'associant à la société de production de vidéos d'information pour les réseaux sociaux Brut en mars 2017[135].

La stratégie de média global que nous venons de décrire permet d'interpréter les positionnements des médias français proposant du contenu sur les enceintes connectées. Nous pensons que la radio a dans ce contexte une opportunité à saisir, car son identité est issue du travail du matériau vocal (c'est son savoir-faire), sans pour l'instant que nous soyons capables de qualifier la manière dont elle se saisit de cet enjeu. Une analyse détaillée de l'offre des médias français sur Amazon Echo — ici choisi à valeur d'exemple — devrait permettre de comprendre le positionnement de la radio sur ce support.

2. Analyse de l'offre des médias présents sur Amazon Echo en France.

Nous considérons la voix comme le matériau du savoir-faire radiophonique, dont les enceintes connectées constituent un aboutissement de ses propriétés d'immersion et d'interaction en contexte de virtualité. Sur ce support, la radio perd cependant le monopole du travail de la voix, accaparé par d'autres acteurs médiatiques régis — radio incluse — par une stratégie de média global qui répond à des enjeux concomitants d'image et de diversification des audiences. L'association de nombreux médias au lancement d'Amazon Echo en France le 13 juin 2018 illustre ce point. Des titres de presse, des télévisions et des radios ont créé leurs *skills* — équivalent des *apps* — pour profiter de l'effet d'annonce suscité par l'arrivée de la technologie du géant américain dans l'hexagone et, inversement, faciliter l'implantation d'Amazon sur le territoire français. Comme nous l'avons mentionné précédemment, les médias sont presque contraints de proposer du contenu sur l'enceinte connectée, dans un mouvement de course à l'innovation : « les médias français ont conscience de ne pas chercher le profit. L'objectif est ailleurs : il y a à la fois l'envie de tester des choses, d'innover de nouvelles manières d'informer, et, forcément aussi, de ne pas prendre le risque d'être à la traîne par rapport aux confrères »[136], résume un article de *Numerama*. Plusieurs médias ont pour ces raisons développé des contenus courts, de deux à cinq minutes, pour Amazon Echo. Nous souhaitons, au sein de cette offre, étudier les discours d'escorte accompagnant le lancement de ces contenus chez les médias partenaires du lancement (qui ne représentent pas la totalité des

---

135 *CB News*, Wojciak, T., « France Télévisions étend son partenariat avec Brut », 07/03/2017. Disponible à l'adresse suivante : http://www.cbnews.fr/digital/france-televisions-etend-son-partenariat-avec-brut-a1033931

136 *Numerama*, Turcan, M., « Comment Amazon Echo s'appuie sur les médias pour conquérir les Français », 06/06/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/380039-comment-amazon-echo-sappuie-sur-les-medias-pour-conquerir-les-francais.html

médias désormais présents), puis focaliser notre attention sur la radio. Nous avons ainsi défini un corpus d'analyse comprenant deux dimensions : les offres des médias partenaires du lancement d'Amazon Echo sur le marché français (a), et l'offre déployée par les stations de radio sur ce support (b).

a) Analyse des discours d'escorte des offres des médias partenaires du lancement d'Amazon Echo en France.

Nous avons défini le corpus pour ce premier temps de notre analyse avec deux critères principaux :

- ***Un contenu est analysé par média partenaire du lancement, et provient d'Internet***. Qu'il s'agisse d'un communiqué de presse ou d'un article comportant la même fonction d'exposition de l'offre (dont les publics, professionnels ou de consommateurs, diffèrent mais reçoivent un message dont la justification de la création de la *skill* a la même teneur), nous avons souhaité limiter notre étude à un contenu pensé comme représentatif du discours légitimant la présence de la marque sur Amazon Echo, et plus détaillé que les courtes descriptions que l'on peut trouver sur la page web regroupant les *skills* d'Amazon en France (amazon.fr/skills).
- **Ce contenu s'inscrit dans des bornes temporelles précises : du 1er au 30 juin 2018**, entourant la date de lancement d'Alexa (le 13 juin 2018) d'environ quinze jours en amont et en aval, afin d'y trouver ce que nous avons qualifié précédemment d'effet d'annonce.

Afin de composer une liste de médias cohérente, nous avons d'abord composé notre corpus grâce à des articles de la presse spécialisée en marketing et en communication regroupant les noms des médias associées au lancement d'Amazon Echo (un de *L'ADN*[137] et un de *CB News*[138]), avant de chercher sur le web des contenus répondant aux critères précédemment énoncés. Nous avons établi une liste de sept contenus, regroupant quatorze médias (le groupe *Les Échos-Le Parisien* rassemblant deux médias et Radio France comptant sept chaînes). Parmi eux, le *Huff Post* —

---

137 *L'ADN*, Huot, A., « Lancement d'Alexa en France : quels médias ont répondu présent ? », 13/06/2018. Disponible à l'adresse suivante : http://www.ladn.eu/media-mutants/les-medias-francais-sur-alexa/

138 *CB News*, Wojciak, T., « Amazon lance Alexa en France : les médias s'y engouffrent… », 06/06/2018. Disponible à l'adresse suivante : http://www.cbnews.fr/mobile/amazon-lance-alexa-en-france-les-medias-s-y-engouffrent-a1043088

anciennement *Huffington Post* — est un site web depuis son origine, ancré dans l'écrit à la fois dans son positionnement éditorial et l'appartenance de sa version française au groupe de presse Le Monde. Le reste des médias représentés dans nos contenus comprend quatre titres de presse écrite (*Les Échos*, *Le Parisien*, *20 Minutes* et *L'Équipe*), une chaîne de télévision (LCI), et huit radios (les chaînes Radio France et NRJ), attestant de l'appropriation du matériau vocal par une pluralité d'acteurs extérieurs à la radio. L'apparente sur-représentation de radios dans notre corpus est en réalité contrebalancée par le fait que le groupe Radio France, en associant franceinfo au lancement d'Amazon Alexa avec sa *skill* de flash info disponible par défaut sur l'appareil, rend ses autres chaînes accessibles dans le même temps sans que ne soient conçus des contenus dédiés pour la plateforme. Voici la liste des sept contenus que nous allons tenter d'analyser, dont les textes sont disponibles en annexe[139] :

- *Le Huff Post*, Phalippou, A., « Le HuffPost se lance dans l'audio sur Amazon Echo », 12/06/2018[140].
- LCI, Ingrand, C., « 'Alexa, quelles sont les nouvelles ?' : LCI vous informe sur l'assistant vocal d'Amazon », 14/06/2018[141].
- *20 Minutes*, Beaudonnet, L., « VIDEO. Alexa en France : Retrouvez toute l'actu de « 20 Minutes » sur Alexa d'Amazon », 22/06/2018[142].
- *Les Échos-Le Parisien*, « Le groupe Les Échos-Le Parisien fait entendre sa voix sur Alexa, l'assistant vocal d'Amazon », 06/06/2018[143].
- NRJ, « NRJ sur Amazon Alexa »[144]. Nous n'avons pu établir la date de création de la page, mais les sources qui nous ont permis d'établir cette liste confirment l'association de la station au lancement d'Amazon Echo. Plutôt que d'exclure NRJ de notre corpus, nous souhaitons apporter cette précision.

---

[139] Annexe 17, pp.214-219.

[140] Disponible à l'adresse suivante : https://www.huffingtonpost.fr/alexandre-phalippou/le-huffpost-se-lance-dans-l-audio-sur-amazon-echo_a_23450332/

[141] Disponible à l'adresse suivante : https://www.lci.fr/high-tech/alexa-quelles-sont-les-nouvelles-lci-vous-accompagne-et-vous-informe-sur-l-assistant-vocal-d-amazon-2090470.html

[142] Disponible à l'adresse suivante : https://www.20minutes.fr/high-tech/2289259-20180622-video-alexa-france-retrouvez-toute-actu-20-minutes-alexa-amazon

[143] Disponible à l'adresse suivante : http://pressroom-lesechos-leparisien.com/2018/06/06/le-groupe-les-echos-le-parisien-fait-entendre-sa-voix-sur-alexa-lassistant-vocal-damazon/

[144] Disponible à l'adresse suivante : http://www.nrj.fr/amazon-alexa

- *L'Équipe*, « L'Équipe lance son flash briefing sur Amazon Echo », 06/06/2018[145].
- Radio France, « Radio France poursuit son développement sur les assistants vocaux avec Amazon Alexa », 13/06/2018[146].

L'objectif de ce regroupement de discours de médias distincts est comparatif : selon leur identité première, les médias mettent-ils en avant les mêmes arguments ? Quels sont les points communs ? Quelles sont au contraire les spécificités des discours de chaque marque ? Au-delà de l'effet d'annonce dont chacune d'entre elles profite, et que les bornes temporelles choisies pour la constitution de notre corpus — quinze jours avant et après le lancement d'Amazon Echo environ — accentuent, nous souhaitons savoir si l'on observe une tendance à l'effacement des spécificités liées au support d'origine pour laisser émerger une marque média, ou si la particularité du support d'origine y est figurée, explicitement ou implicitement. La finalité de cette étude comparative est d'évaluer la prégnance la notion de média global dans les stratégies médiatiques à l'aune du lancement d'Amazon Echo. Commençons par étudier nos contenus séparément avant d'établir une comparaison rassemblant ces notions.

Le HuffPost « accompagne la sortie de l'assistant vocal d'Amazon avec deux nouveaux formats d'émissions en audio », constituant chacun une *skill* distincte. Avant de détailler leurs contenus, le *pureplayer* met leur création en contexte, rappelant le besoin de pédagogie à propos des enceintes connectées que nous avons préalablement identifié chez le public français. La nouveauté de l'objet est ainsi soulignée par le biais de sa singularité — « ces drôles de petits cylindres, palais ou boules appelés 'assistants connectés' qui veulent envahir votre domicile » —, pour finalement en préciser l'utilité : « Une enceinte sans fil qui fait également radio, rien de bien neuf sous le soleil ? Sauf que les assistants vocaux comprennent ce que vous dîtes et ça, ça change tout ». Qu'est ce qui change ? La possibilité d'une interaction directe entre l'utilisateur et le contenu, qui est vécue comme une révolution des usages — elle « pourrait bien modifier les comportements dans la façon de s'informer » — et de la production médiatique, parlant de l'instauration d'une « écriture nouvelle ». En tant que média global, le *HuffPost* se met en scène comme un média ancré dans l'innovation — « C'est pourquoi le *HuffPost*, fidèle à son credo d'informer sur l'actualité en restant innovant et surprenant, se lance aujourd'hui dans l'audio » —, et qui souhaite se distinguer de ses confrères sur le support des enceintes connectées. Il est intéressant de noter que la radio est citée

[145] Disponible à l'adresse suivante : https://blog.lequipe.fr/operations/lequipe-amazon-echo/

[146] Disponible à l'adresse suivante : http://www.radiofrance.fr/espace-pro/espace-presse/communiques-de-presse/2018/06/radio-france-poursuit-son-developpement-sur

spontanément, confirmant le lien de son identité au travail de la matière sonore et vocale : « avec l'envie de faire des émissions qui ne ressemblent pas, sur la forme, à de la radio et offrent des informations supplémentaires de ce qu'offrent les grandes stations ». Le savoir-faire radiophonique dans la production d'information audio est ici à tel point confirmé que le *HuffPost* semble ne pas souhaiter le concurrencer sur ce terrain, empruntant la tangente des contenus originaux et décalés (des informations « supplémentaires »). Parmi les deux formats proposés, l'un d'information et l'autre de coaching, le premier est celui qui concentre notre intérêt car il résonne avec les usages des enceintes connectées que nous avons identifiés, portés sur l'actualité notamment le matin. « Quoi de neuf le Huff ? » consiste en une mise en abyme, où le titre d'information se dote de sa propre voix artificielle, le robot Huffie, avec laquelle un journaliste dialogue sous la forme de questions-réponses, à la manière d'un utilisateur avec son enceinte connectée. Une information « surprenante », « amusante, choquante, décalée ou simplement passée inaperçue » est traitée chaque jour.

LCI spécifie d'emblée son identité de média télévisé. Avec sa *skill*, la chaîne d'information en continu proposera chaque jour un « Essentiel de l'Info » à la fois audio et vidéo — « un résumé de l'actualité, toujours à jour, et disponible à la demande, en son et en images » —, la version la plus grande de l'Amazon Echo — le Echo Spot — étant dotée d'un écran LCD. L'origine télévisuelle est mise en avant jusque dans le vocabulaire employé, son « Essentiel de l'Info » consistant en un « JT » (Journal Télévisé). Comme le *HuffPost*, le texte de présentation de LCI contient une entrée en matière à portée pédagogique, captures d'écran à l'appui, sur la manière d'installer sa *skill*. Souhaitant créer un réflexe vers son média, LCI s'associe à une situation d'écoute proposée par les concepteurs d'Amazon Echo : le flash quotidien, que l'utilisateur peut personnaliser en installant les *skills* dont il souhaite entendre les contenus chaque jour, et leur ordre d'écoute. Quand le *HuffPost* insiste sur l'interactivité que les enceintes connectées permettent entre l'utilisateur et le contenu, LCI focalise l'attention du lecteur sur ses facultés de personnalisation : « Démarrer par un flash info, enchaîner sur la météo, écouter son horoscope, puis son podcast préféré, c'est l'utilisateur qui décide du contenu, et du moment où il l'écoutera ».

20 Minutes mêle dans la présentation de son offre ces dimensions d'interaction et de personnalisation avec une vidéo mettant en scène plusieurs situations d'écoute (dans son salon en journée, dans sa salle de bain le matin, ou lors d'une discussion en cuisine où l'on a besoin de faire une recherche sur Internet pour vérifier la validité d'un sujet abordé au cours de la conversation). L'élément de cette vidéo présentant le plus d'intérêt se situe à la fin, lorsque l'on voit le journaliste auteur d'un article en travailler l'écriture. *20 Minutes*, en tant que titre de presse, a en effet décidé

de proposer la lecture de ses articles en *text-to-speech* (l'Intelligence Artificielle Alexa lit les articles), au résultat désincarné et présentant aujourd'hui des défauts de fluidité dans l'enchaînement des phrases. En montrant le journaliste à l'œuvre, l'origine humaine du contenu est rappelée. L'utilisateur peut accéder à un flash info regroupant les trois articles essentiels de la journée écrits par la rédaction de *20 Minutes*, et demander la lecture de l'article le plus populaire (sans que les critères de définition de sa popularité ne soient précisés) d'une rubrique spécifique. Si le choix du *text-to-speech* n'est pas justifié dans le contenu de présentation que nous étudions, il est d'abord d'ordre économique. Il permet néanmoins de figurer une identité de média papier, qui s'exprime par son écriture ici convertie en lecture — « [Alexa] vous dira tout haut ce que, d'habitude, vous lisez tout bas ». Les contenus les plus concrets, à caractère purement informatif, sont lus par l'Intelligence Artificielle. Pour écouter un contenu au « ton décalé », l'utilisateur doit se diriger vers la « Minute Papillon », un podcast conçu et incarné par une vive voix — plus précisément, « la douce voix de notre journaliste Anne-Laetitia Béraud qui vous donne les informations principales de la journée sur un ton décalé ». Au sein même de la *skill* de *20 Minutes* se déploient deux axes : celui de l'originalité et de l'intimité, avec une voix humaine, et celui plus pragmatique de la voix artificielle. Une dose d'humanité semble nécessaire pour incarner le contenu, de la même manière que Huffie, le robot du *HuffPost*, ne peut parler seul à l'utilisateur. Il lui faut dialoguer avec un journaliste.

Le groupe Les Échos-Le Parisien présente dans le contenu de présentation que nous étudions deux offres distinctes, chacune incarnée par l'un de ses titres — le « Flash Audio » du *Parisien* et « Les Éditos de la Rédaction » des *Échos*. Le texte se structure en trois parties : une mise en contexte, puis le détail de l'offre du *Parisien* et enfin de celle des *Échos*. L'entrée en matière revêt à nouveau un caractère pédagogique mêlant l'utilité de l'objet et ses situations d'écoute : « À la maison comme au bureau, Alexa est conçue pour simplifier la vie de ses propriétaires en leur permettant de contrôler leur quotidien grâce à leur simple voix ». À l'inverse de *20 Minutes*, le groupe a fait le choix du podcast (dans un format de trois minutes), dépréciant la qualité du *text-to-speech*. Pierre Chausse, directeur des rédactions du *Parisien*, le présente ainsi : « Nous avons imaginé un produit très différent d'une simple version « text-to-speech » mais également de ce que l'on peut entendre en radio ». En effet, le *Parisien* souhaite figurer son ancrage dans l'écrit en incarnant ses contenus de vives voix. Ces dernières sont immédiatement audibles dans le contenu proposé (« le ton employé »), et représentées par le corps collectif des journalistes, qui en assurent l'originalité éditoriale (« (…) notre flash est composé de quatre à cinq sujets forts choisis par la rédaction »). La voix doit être vive, à la manière d'un style d'écriture qui émane de la

personnalité de son auteur, pour retranscrire « tout ce qui constitue l'identité du *Parisien* et en fait un média d'excellence à l'écrit ». Le choix de la voix humaine suppose un investissement financier supérieur à celui du *text-to-speech*, et d'un savoir-faire encore balbutiant au sein d'un média d'origine papier qui n'avait pas de savoir-faire dans le domaine de la production audio. Les équipes du journal ont ainsi fait appel à l'expertise du studio de podcasts natifs français Binge Audio, qui offre lui-même des contenus sur Amazon Alexa, comme le rapporte un article de *Numerama*[147]. *Les Échos* proposent quant à eux l'un des trois podcasts qu'ils produisent déjà sur l'enceinte connectée d'Amazon. « L'Édito de la Rédaction » propose d'écouter chaque jour, « l'analyse et l'opinion de l'un de nos éditorialistes sur le sujet à la Une de l'actu éco ». En tant que média spécialisé dans l'actualité économique, *Les Échos* incarnent cette expertise avec la vive voix de l'un de leurs éditorialistes, plutôt que de le reléguer au statut d'auteur d'un contenu lu par Alexa. « L'analyse et l'opinion » demeurent des facultés humaines dont les modalités de la *skill* du journal tentent de préserver la qualité, afin de ne pas diluer leur identité de média dans ce déploiement opéré sur les enceintes connectées, à l'empreinte artificielle. Contrairement à *20 Minutes*, le groupe Les Échos-Le Parisien fait le choix de la voix humaine comme gage de qualité, et vécu comme tel par les utilisateurs. Un regard rapide aux commentaires clients sur la page web de la *skill* sur le site d'Amazon[148] le confirme : « Merci Le Parisien de ne pas avoir cédé à la voix automatique d'Alexa ! Flash très agréable à écouter, dont le contenu est différent de ce que l'on entend ailleurs. J'apprends des choses tous les matins. Je recommande » ; « Agréable à écouter ! Intéressant de découvrir le meilleur du journal sur l'enceinte. Il offre une véritable fraîcheur par rapport aux flash des autres médias. Voix agréables. Je recommande ! ». La radio, qui manie historiquement le matériau vocal à des fins d'information, est par conséquent dotée d'une avance sur les titres de presse qui souhaitent user du même type de médiation. Elle possède déjà ce savoir-faire, et l'oralité est constitutive de son identité de média.

NRJ, radio musicale, est un cas à part dans notre analyse. Contrairement aux autres éléments de notre corpus, ses concurrents les plus puissants sur le numérique ne sont pas les autres stations ou médias, mais les plateformes de *streaming* musical telles que Spotify ou Deezer — constat que Julien Le Ny, producteur chez Fun Radio avec qui nous nous sommes entretenus,

---

147 *Numerama*, Turcan, M., « Comment Amazon Echo s'appuie sur les médias pour conquérir les Français », 06/06/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/380039-comment-amazon-echo-sappuie-sur-les-medias-pour-conquerir-les-francais.html

148 Disponible à l'adresse suivante : https://www.amazon.fr/Le-Parisien/dp/B07CJTGG8S/ref=sr_1_1?s=digital-skills&ie=UTF8&qid=1535216145&sr=1-1&keywords=le+parisien

confirme en faisant également référence aux quotas de musiques imposés par le CSA à chaque station :

> « Malheureusement la radio est un média qui vieillit et qui a du mal à évoluer parce qu'il y a beaucoup de règles qui le régissent. On n'arrive pas à pouvoir s'aligner sur nos concurrents, qui ne sont pas obligatoirement les autres radios mais plutôt les Deezer ou Spotify qui peuvent diffuser ce qu'ils veulent, peuvent s'adapter personnellement à l'auditeur... Chose qu'on ne peut pas faire. On n'est pas sur du "sur-mesure", on est sur un format musical et un cœur de cible qui doit être satisfait quand il écoute notre radio »[149].

NRJ met ainsi en avant, au sein d'une page davantage visuelle que rédigée, sa fonction de programmation et d'éditorialisation, que les plateformes concurrencent grâce aux capacités de personnalisation développées par leurs algorithmes. NRJ propose d'écouter son flux en direct, et d'accéder à diverses *playlists* composées de « hits » du moment, conformément à son positionnement en FM, divisés en différents genres de musiques.

L'Équipe présente une offre moins développée sur Amazon Alexa que sur Google Home, où elle ajoute à son flash un « Trivia Game », où l'utilisateur peut interagir avec son enceinte connectée dans le cadre d'un quizz sur une thématique sportive, et une « Histoire du Jour ». Le « Flash l'Équipe » est introduit par une mise en contexte pédagogique, commune à l'ensemble des éléments de corpus que nous avons étudiés jusqu'à présent, avant une succincte description du flash : « un court résumé des faits marquants de la veille et des principales échéances de la journée à venir », dans le domaine de l'actualité sportive. Très synthétique, ce contenu ancre son média dans une démarche d'innovation portée par un enjeu sous-jacent d'image, et par la volonté de diversifier ses audiences. Emmanuel Alix, directeur du pôle numérique du journal, est cité : « (…) Cette collaboration avec Amazon s'inscrit dans le prolongement d'une stratégie d'innovation qui vise à proposer le meilleur contenu possible à nos clients, quels que soient leurs usages de consommation » — le média propose par ailleurs des podcasts.

Radio France publie le communiqué de presse que nous étudions dans un contexte particulier. Sa chaîne d'information en continue, franceinfo, est partenaire d'Amazon lors du lancement de son enceinte connectée en France (et y est donc plus attachée qu'un média se contentant de s'y associer). Elle est grâce à cela installée par défaut sur les appareils de la gamme

[149] Entretien avec Julien Le Ny, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 2, pp.148-153.

Echo, répondant à la question « Alexa, quelles sont les dernières infos ? »[150], et fait l'objet d'un paragraphe à part dans le communiqué que nous étudions. Le groupe Radio France profite de cette collaboration pour proposer l'écoute spécifique de plusieurs de ses podcasts, en plus de l'accès aux flux en direct de ses sept chaînes: l'émission « C'est déjà demain » d'Hélène Chevalier et « Géopolitique » de Bernard Guetta sur France Inter (qui ne poursuit pas sa chronique à la rentrée 2018-2019 sur la station), ainsi que le « Billet Culturel » de Mathilde Serrell sur France Culture. Les composantes du flash de franceinfo dédié à Amazon Alexa illustrent sa volonté de se figurer comme un média global conservant une identité liée au travail de la voix. Deux versions sont en effet proposées : un flash audio réactualisé quatre fois par heure, et une version vidéo réactualisée trois fois par heure, ces flashes provenant de sa station de radio et de sa chaîne de télévision. Le média public conjugue par conséquent la figuration de son identité de média et l'implantation par défaut dans une enceinte connectée, lui conférant un avantage concurrentiel sur les autres titres d'actualité. D'autres *skills* sont proposées par franceinfo : un point sur l'économie dédié à la plateforme ainsi que des rattrapages de plusieurs de ses émissions, dont nous proposerons un détail dans notre analyse des offres des radios sur l'enceinte connectée d'Amazon. Le matériau vocal — qui, étendu au son, constitue l'audio — semble être pensé par le groupe Radio France comme une substance qui se déplace de support en plateforme selon les usages, érigeant le nom de son site spécialisé dans la veille du marché audio numérique, « Hyperradio », au rang de méta-code de sa stratégie de média global. Son développement sur les enceintes connectées répond en effet à l'enjeu dessiné dans la note de Mathieu Gallet pour Terra Nova que nous avons citée, le communiqué précisant que son association à Amazon Echo « s'inscrit dans la stratégie de développement de Radio France face à la révolution des usages », « pour répondre à la démultiplication des supports d'écoute audio », le groupe s'attachant « à rendre ses radios hyper accessibles en cherchant à produire, éditer, et diffuser le bon contenu, au bon moment et sur la bonne plateforme ».

Que conclure de cette analyse ? Les médias que nous avons répertoriés et dont nous avons étudié les discours d'escorte de leurs offres sur Amazon Alexa s'inscrivent tous dans une logique de média global, qui multiplie les supports de sa prise de parole tout en employant un dosage entre nouvelle expression et figuration d'une identité de marque média liée à son origine. Chacun met par ailleurs son innovation en contexte, confirmant le manque de compréhension du public français

---

[150] *Franceinfo*, « franceinfo, n°1 de l'info en continu sur les assistants vocaux, disponible via les services d'Amazon », 13/06/2018. Disponible à l'adresse suivante : https://www.francetvinfo.fr/internet/amazon/franceinfo-n1-de-l-info-en-continu-sur-les-assistants-vocaux-disponible-via-les-services-d-amazon_2798817.html

quant à l'utilité et aux fonctionnalités des enceintes connectées — ou peut-être, plus simplement, sa nouveauté. Plus largement, nous observons que les marques médias qui ne proviennent pas de la radio doivent faire un effort supplémentaire pour légitimer leur présence dans cet environnement vocal — un angle « décalé » pour le *HuffPost*, du *text-to-speech* pour *20 Minutes*, le recours au savoir-faire de Binge Audio pour le groupe Les Échos-Le Parisien, par exemple. L'offre du groupe public Radio France avec franceinfo apparaît comme la plus développée, et ce pour deux raisons : son ancrage dans la matière vocale est figuré par l'utilisation des voix de sa station dans ses flashes, et sa collaboration avec Amazon facilite l'accès à son service. Les entreprises de presse sont en revanche celles qui font montre du plus d'originalité, créant des contenus dédiés à la plateforme. Les médias thématiques, allant de la radio musicale NRJ au quotidien sportif *L' Équipe*, ont des statuts éminemment liés à leurs propositions éditoriales, ne faisant pas entrer en jeu les mêmes types de concurrences et de productions. Afin de comprendre la manière dont le média radio se positionne au sein de ce champ concurrentiel élargi, nous allons maintenant étudier l'offre radiophonique sur Amazon Alexa.

b) Analyse de l'offre des stations de radio françaises sur Amazon Alexa.

Un tableau présent en annexe[151] propose une photographie de l'offre des stations de radio françaises sur Amazon Echo au 12 août 2018. Elle consiste en un répertoire des *skills* de stations de radio présentes sur la bande FM (et de webradios leur étant directement associées, comme la radio musicale Mouv' du groupe Radio France) apparaissant à la requête « radio » sur amazon.fr/skills. Notre analyse ne se resserre désormais plus à la période du lancement de l'enceinte connectée, en juin 2018, mais à l'ensemble des contenus proposés à l'été. Ces derniers seront probablement complétés d'autres propositions et de l'apparition de nouveaux acteurs dans les mois suivants. Notre objectif est de qualifier les types de contenus proposés par les radios selon trois critères :

- **Le flux en direct** est-il proposé sur Amazon Echo ? Ce mode de diffusion, qui définit la radio historiquement, répond à la fonction d'accompagnement de ce média et de l'enceinte connectée, présente à chaque étape du quotidien. Il consiste autant en une figuration de l'identité de la radio qu'en l'absence d'une proposition dédiée à l'enceinte connectée.

[151] Annexe 18, pp.220-222.

- **Des contenus délinéarisés**, qu'ils s'agissent de podcasts proposés sur d'autres plateformes ou de contenus plus courts, dédiés à l'enceinte connectée sont-ils inclus dans la *skill* ? Ce mode de diffusion, qui n'est pas l'apanage de la radio, permet tout de même de figurer un ancrage dans l'audio qui serait multi-plateformes, affirmant la présence de ce média dit « classique » sur le numérique.
- **Des contenus produits uniquement pour l'enceinte** sont-ils présents dans l'offre des radios sur Amazon Echo ? Nous avons conclu de nos précédentes observations que la logique structurant la notion de média global supposerait que ce soit le cas. La présence de contenus dédiés à l'enceinte — à l'instar des contenus des titres de presse tels que le *HuffPost* ou *L'Équipe* — attesteraient d'un effort des stations de radio pour proposer des modes de consommation de leurs médias adaptés aux usages de l'enceinte connectée au plus haut degré.

Notre tableau répond par l'affirmative (« oui », ce type de contenu est présent) ou la négative (« non », ce type de contenu n'est pas présent) à ces questions, nous apportant des données quantitatives sur l'offre des radios sur les enceintes connectées. Les vingt-deux stations[152] identifiées proposent un total de trente-cinq *skills*, certaines proposant plusieurs *skills* pour l'écoute délinéarisée de différents programmes de leurs grilles, répartis en thèmes. France Bleu propose trois *skills* (« France Bleu », « France Bleu Santé & Bien-Être », « France Bleu Sciences »), France Culture deux (« France Culture », « France Culture - Actualités Culturelles »), franceinfo cinq (« Franceinfo », « Le Flash Franceinfo » — présent par défaut sur l'enceinte —, « Franceinfo - hightech », « Franceinfo - économie », « Franceinfo - sports »), France Inter quatre (« France Inter », « France Inter - Humour », « France Inter - Géopolitique », « France Inter - Hightech »), et RFI trois (« RFI », « RFI Monde », « RFI - Journal en Français Facile »). Nous constatons qu'aucune radio privée présente dans notre corpus ne propose plus d'une *skill*, attestant peut-être d'un désir de diversification de l'audience moins attaché à des problématiques économiques chez le groupe public Radio France, sans que nous puissions confirmer cette hypothèse. Nous n'échappons pas au possible mauvais référencement de certaines *skills*, dans ce cas problématique pour les radios concernées. Les résultats de notre analyse, effectués sur un total de 35 *skills*, sont les suivants :

---

[152] Dans l'ordre des résultats : Sud Radio, Fun Radio, France Bleu, Virgin Radio, France Musique, Mouv', RTL2, FIP, France Culture, Franceinfo, France Inter, Chérie, Nova, NRJ, RFM, Europe 1, Skyrock, RTL, RFI, RMC, Rires et Chansons, Les Indés Radios.

- **Dix-neuf *skills* contiennent le flux en direct de leur station**, et seize ne le proposent pas. Cette donnée doit être nuancée par l'existence de plusieurs *skills* pour certaines stations. Si nous ramenons la proposition du flux en direct non pas au nombre de *skills* mais au nombre de marques médias, nous observons que seules trois stations ne proposent pas leurs flux en direct : Radio Nova, RMC et Rires & Chansons.
- **Vingt-huit *skills* proposent des contenus délinéarisés**, contre seulement sept qui en sont dénuées. Ce chiffre peut soit signifier l'absence d'offre numérique très développée, soit le choix de ne proposer que le flux en direct. A part Sud Radio, nous constatons que seules des radios musicales ne proposent pas de contenus délinéarisés, la « Revue de Presse Musicale de Nova » faisant office d'exception. Les programmes de *talk* étant moins présents dans leurs grilles, la proposition d'une offre de musique délinéarisée pourrait faire perdre à ces radios leur fonction de programmation en direct, qui les distingue des plateformes de *streaming* musical telles de Spotify ou Deezer. Le contenu proposé par Radio Nova en délinéarisé est parlé, et initialement proposé aux auditeurs de sa matinale.
- **Aucune station de radio ne propose de contenu créé uniquement pour Amazon Alexa**, les offres délinéarisées consistant toutes en la récupération de séquences diffusées en direct à l'antenne. Pour exemple, le « Flash Franceinfo » que nous avons mentionné précédemment est un enregistrement du *flash* proposé plusieurs fois par heure sur la bande FM. Si le chiffre zéro interpelle, il peut correspondre à deux problématiques : le souhait de figurer son identité de média radio et d'amener les utilisateurs d'Amazon Echo à écouter le direct avec un contenu d'appel, ou un souci d'économie d'échelle, multipliant les supports de diffusion d'un même contenu à moindre coût dans un contexte de faible développement du marché des enceintes connectées en France. Nos précédentes considérations ont cependant identifié la pertinence de distinguer son offre numérique de son offre en direct, afin d'affirmer sa légitimité sur chaque support de prise de parole.

À l'été 2018, l'offre des stations de radio françaises sur Amazon Echo fait montre de peu d'originalité. Si l'usage de la voix dans une situation d'interaction directe semble être un enjeu pour les antennes, comme notre analyse des discours d'escorte des marques médias présentes sur l'enceinte a semblé le montrer, les ventes de l'objet demeurent trop confidentielles pour que ce champ de création soit pleinement investi. La radio a un avantage concurrentiel sur ce support lié à son ancrage dans le travail de la voix, mais elles ne sont aujourd'hui pas saisies des possibilités narratives déployées par l'interaction d'un utilisateur avec le contenu. Sans prononcer d'avis, nous

remarquons que les titres de presse font preuve de davantage d'inventivité sur ce support, car elles se déplacent dans un univers qui ne les concerne pas initialement, les obligeant à se distinguer de la radio. Cette dernière se trouve dans une position ambivalente, où le juste dosage d'une figuration d'une identité sonore et vocale entre en conflit avec les usages proposés sur les plateformes numériques, dont les enceintes connectées constituent un aboutissement des logiques d'interaction et d'immersion.

En ce temps de réflexion pour la radio, d'autres médias se saisissent de son savoir-faire. Il nous faut à présent étudier le deuxième aspect qui contraint ce média sur le support des enceintes connectées : le déplacement de la voix incarnée par un corps vers les timbres robotiques d'Intelligence Artificielles faisant l'objet d'investissements colossaux de la part de leurs concepteurs. Si ces « intelligences » sont déjà à l'œuvre dans nombre de supports numériques — pour exemple les algorithmes des réseaux sociaux ou des plateformes de *streaming* —, les enceintes connectées sont une étape supplémentaire du déplacement et de l'encadrement de la production médiatique par la machine, puisqu'elle se saisissent d'un matériau défini par son référent sensible, sa preuve du vivant qui l'incarne. Analyser les modes de conception de ces voix artificielles (qui se font toujours dans un souhait d'adoption de la part des publics) en rapport avec la voix de radio pourrait selon nous esquisser les contours du futur des médias sur le numérique, où la création humaine sera de plus en plus accaparée par des Intelligences Artificielles permettant à leurs créateurs d'affirmer leur hégémonie sur des marques médias qui répondent déjà en de nombreux points aux règles qu'elles imposent.

C. DE LA RADIO À L'ASSISTANT VOCAL : LA DÉSINCARNATION DE LA VOIX COMME ENJEU TECHNOLOGIQUE.

Nous avons, depuis le début de notre recherche, suivi le voyage de la voix dans une pluralité d'espaces — l'extériorisation du sujet dans le monde, le transport de son corps dans les ondes radiophoniques, et le déplacement de sa source vers une machine. Si nous avons étudié les modifications de la relation que la voix instaure dans ces espaces, allant de l'effet d'authenticité du média radio pour son auditeur à la proposition d'interaction des enceintes connectées, déployant un champ concurrentiel élargi par la notion de média global, nous n'avons abordé qu'une partie de son

altération. Un aspect central reste en effet à étudier : les implications technologiques et sensorielles d'une voix dont la source n'est plus un homme, mais une machine. La voix est dans ce mouvement désincarnée, elle perd le corps qui l'origine et par-là même son *grain*, trace de son référent sensible. Nous allons analyser les deux aspects concomitants qui régissent le déplacement de la voix d'une source naturelle vers une source artificielle : l'altération du corps par la technologie, indissociable du travail médiatique, qui construit un espace dans lequel l'homme est dématérialisé (1), et les logiques de conception de cette voix artificielle chez ses créateurs, commandées par la volonté d'effacer la trace non plus d'un transport altérant les propriétés de la vive voix, mais d'un déplacement de sa source (2). La place du corps, de l'humain dans ce contexte, sera le fil conducteur de nos réflexions, car elle contribue selon nous à définir le statut de la radio sur les enceintes connectées, et plus largement au sein de l'économie numérique.

1. L'altération de la vive voix dans son transport technologique : un processus déjà effectif en radio.

La voix, en tant que conjugaison des sonorités de nos tempéraments et d'un arbitraire du signe linguistique, incarne le déploiement d'une personnalité vers une altérité, celle du monde physique ou des autres membres de l'espèce humaine. « Faire entendre sa voix, babiller, parler, chanter, rire ou pleurer, c'est vivre en sujet dans le monde des hommes »[153]. Le transport de la voix dans un espace médiatique — en un mot, créer une médiation — instaure une distance, qui en altère les composantes émotionnelles, supposant un travail pour en restituer l'authenticité. Ce travail, nous en avons évoqué une partie avec les voix de radio. Il s'étend pourtant à toutes les narrations médiatiques. Les travaux de Luc Boltanski, sociologue et directeur d'études à l'École des Hautes Études en Sciences Sociales (EHESS), sur le mécanisme de la compassion[154], qui ne peut se réaliser naturellement que dans une situation d'immédiateté avec la scène de souffrance source de sympathie, et suppose d'en restituer la proximité à des publics qui en sont éloignés lorsqu'ils en visionnent ou en écoutent le reportage télévisé ou radiophonique, sont un exemple. Dès que la voix n'est plus vive au sens strict du terme, c'est-à-dire que nous interagissons directement avec sa source, elle est altérée.

Ce phénomène est à l'œuvre en radio. Le travail radiophonique de l'oralité consiste à effacer la trace de son maniement de la matière vocale et sonore au cours du processus de création pour

---

153 Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004, p.11.

154 Boltanski, L., *La Souffrance à Distance. Morale Humaine, Médias et Politique*, éd. Métailié, 1993.

créer son effet d'authenticité. Le surgissement d'une erreur est maladroit (pour reprendre les termes d'Erving Goffman[155]) car il rappelle cette distance, ici réifiée par la rupture du contrat qui unit l'animateur radio à ses auditeurs. Ce travail de la distance insérée dans la voix lors de son transport technologique est présent dans plusieurs dimensions de la radio, qu'il s'agisse du discours de l'animateur que nous avons évoqué dans une précédente partie ou de la polyphonie d'une émission, incluant des allers-retours entre le studio et l'extérieur par le biais de duplex ou d'interventions téléphoniques. Patrick Charaudeau résume cette polyphonie, régie par la notion de maniement de la distance introduite par le transport technologique de la voix, de la manière suivante :

> « La multiplication des conditions de prise de parole est due au développement des technologies de la communication. A l'émission en direct du studio, s'est ajoutée l'émission enregistrée, en différé, surtout depuis l'apparition du magnétophone au lendemain de la deuxième guerre mondiale ; le duplex permet de réunir dans la même émission des paroles prononcées dans des lieux éloignés de plusieurs milliers de kilomètres ; le couple radio-téléphone fut utilisé dans les émissions de jeux et ensuite dans d'autres types d'émission, il favorisa l'aspect communicationnel de la radio, qu'il s'agisse des émissions de Ménie Grégoire ou des émissions de nuit »[156].

Outre cette dimension énonciative qui se réalise explicitement et en direct dans un programme, l'incorporation d'une distance est effective dès le travail de la matière sonore par les travailleurs du son, désignant par ce terme les réalisateurs, les régisseurs-monteurs et toute autre profession de radio maniant le son avec des outils techniques. Cet aspect est analysé par Pierre Schaeffer, qui a entre autres activités fondé le service de recherche de l'ORTF (Office de Radiodiffusion-Télévision Française) dans les années 1960, avec la notion de « dynéma ». Nous nous fions pour retranscrire cette théorie au texte de Fabien Le Tinnier[157], de l'université de Lausanne, qui explique la théorie de Schaeffer en rassemblant trois textes de l'auteur (son *Essai sur la radio et le cinéma*, la retranscription de l'une de ses conférences intitulée « Le Laboureur et ses Enfants », et « L'Art et les Machines ») pour plus de clarté.

Le chercheur entreprend de qualifier les « transformations » que subit la voix « entre la prise de son et sa diffusion, d'une façon précise entre la diffusion initiale fournie par le microphone et la

---

155 Goffman, E., « Radio Talk », in *Forms of Talk*, éd. University of Pennsylvania Press, 1981, p.320.

156 Charaudeau, P., « Introduction : Problèmes d'Analyse des Médias », in Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984, p.13.

157 Le Tinnier, F., « Dynamique de la voix enregistrée selon Pierre Schaeffer (1941-1963) », in *Entrelacs*, 2014. Disponible à l'adresse suivante : https://journals.openedition.org/entrelacs/621

modulation finale envoyée à l'émetteur [ici l'objet technique] ». Le « dynéma » est l'équivalent du cinéma pour ce qu'il nomme les « objets sonores » (rassemblant la voix, la musique et les bruits), au sein duquel le matériau vocal est transformé à chacune des étapes qui le composent (allant de son enregistrement à sa diffusion à l'auditeur). Pierre Schaeffer distingue deux transformations simultanées des objets sonores : l'une « intensive » et l'autre « de caractère ». La première désigne les modulations de hauteur de fréquence sonore et de profondeur, notamment exprimées en termes de focalisation et de volume, de la voix. La seconde transformation est « statique » lorsqu'elle concerne une émission diffusée en direct, sur laquelle les modulations du technicien ne peuvent être que minimes ; elle est « dynamique » lorsque l'émission est enregistrée, car les modulations potentiellement réalisables sont bien plus nombreuses. C'est pour ces émissions enregistrées qu'entre en jeu la notion de « dynéma », car le « caractère dynamique, lui, non seulement autorise mais requiert généralement le montage, le mixage et l'art de l'ellipse, à la manière d'un monteur de cinéma qui colle et raccorde les plans entre eux pour ainsi constituer l'œuvre d'art »[158]. Dans le cadre d'un contenu « dynamique » sur lequel le technicien peut déployer la totalité de son art radiophonique, cinq types de transformations de la voix, émanation du corps, sont effectuées : la prise de son avec le microphone, l'enregistrement avec les technologies de phonographie, le mixage et le montage, la diffusion, puis la réception. Le mixage et le montage constituent ce que Schaeffer nomme le « dynamophone », qui constitue l'altération majeure de la voix. Le technicien privilégie en effet les éléments de sens « au détriment du reste des composantes de la voix exprimée (ex : le bruit alentour, la musicalité de l'élocution, le grain, la hauteur, *etc.*) »[159]. Le message ainsi créé est d'un « genre nouveau »[160], qui n'appartient plus à la vive voix.

Le travail de la voix en radio, au-delà de ses aspects énonciatifs, a par conséquent une dimension technique qui en altère les qualités puisque sont privilégiés les éléments de sens linguistiques aux sonorités dont le référent est sensible, et qui communiquent elles aussi un message (qu'Iván Fónagy interprète comme un second encodage du code langagier). Comme l'indique le type de transformation que désigne Schaeffer, le « caractère » de la voix de radio n'est plus celui de la vive voix. Cette altération ne doit pas être entendue sous un angle qualitatif. Nous l'avons vu, la radio produit de nombreux effets chez ses auditeurs (d'intimité, de communauté, de réel et d'imaginaire ou d'évocation) que nous avons rassemblés sous le terme d'authenticité, accentués par

158 ibid.

159 ibid.

160 ibid.

un dispositif acousmatique. Il s'agit davantage d'un changement d'espace d'expression pour la voix, qui en transforme les propriétés. La voix de l'animateur se déploie dans un contexte radiophonique, tout comme la voix d'un chanteur enregistrée sur un CD sera enveloppée d'un certain travail du son en studio, qui lui confère des effets particuliers, que nous pensons comme une couche de sens supplémentaire. Pour illustrer notre propos, nous citerons Greil Marcus, critique musical américain, décrivant un enregistrement de « Last Kind Word Blues » chanté par Geeshie Wiley[161]. Les forts grésillements qui accompagnent le chant sur la bande, provenant de l'état des techniques d'enregistrement des années 1930, créent selon lui un langage à part entière — un « contexte sonore », condition et réalisation de l'œuvre :

> « "Last Kind Word Blues", comme l'a dit un auteur, "repose sur une couche de grésillements". Mais ce n'est pas un simple bruit de fond. La différence réside dans le fait que ces grésillements, le murmure de la rivière Milwaukee, sont désormais inséparables de la voix qui chante. Ils nimbent l'interprétation d'une aura de lointain, d'égarement, d'abandon qui tout entière lui appartient depuis le départ. Le bruit de fond — "qui donne l'impression que la chanson peut disparaître à tout moment" — génère un langage au sein duquel même le mot le plus clairement prononcé demeure ambigu dans son contexte sonore, et ce contexte lui-même n'est perceptible qu'à travers ce voile »[162].

2. La voix artificielle réduite à un déterminisme technologique : la quête d'une qualité naturelle dans les logiques de conception des enceintes connectées.

La singularité de la voix artificielle des enceintes connectées ne provient pas de son altération du matériau vocal dans un espace virtuel puisque, comme nous venons de l'observer, tout transport technologique de la vive voix engendre son altération. C'est parce que la source de la voix se déplace d'un corps à une machine que ses propriétés sont transformées, et ce radicalement. En termes de qualité, il suffit de penser à un article en *text-to-speech*, ou plus généralement aux tons monotones d'Amazon Alexa, Google Assistant et SIRI pour s'en rendre compte. Le caractère artificiel de la voix émanant des enceintes connectées est en cela intimement lié à un déterminisme technologique. Les efforts de leurs concepteurs finiront probablement par produire une voix artificielle à résonance naturelle. Deux critères sont ainsi travaillés : la fluidité d'expression de la

---

161 L'extrait est disponible sur YouTube, « Geeshie Wiley - Last Kind Word Blues », à l'adresse suivante : https://youtu.be/oAKfy2W70Qg

162 Marcus, G., *Three Songs, Three Singers, Three Nations*, trad. de l'Américain par Godard, G., éd. Allia, 2018, pp. 64-65.

voix en tant que matériau sonore qui permet l'immersion de son utilisateur dans un contexte où les mondes virtuel et physique se prolongent l'un l'autre ; la capacité de compréhension et de conversation de la voix comme incarnation d'une machine qui entend faire oublier son origine artificielle grâce à la qualité naturelle de son interaction avec l'utilisateur. Avant de traiter ces deux dimensions (b), nous souhaitons nous attarder sur ce que ce déplacement d'une source naturelle à source artificielle de la voix engendre chez les concepteurs et les utilisateurs, dont l'adoption complète de la technologie est l'objectif (a).

a) L'objet intelligent, une notion antithétique rassemblée par l'idéologie du progrès technique.

Lorsqu'un utilisateur dialogue avec son enceinte connectée, « l'interface vient (…) bouleverser l'humanisme classique, déniant à l'être humain son exceptionnalité dans un domaine qui lui était jusqu'alors exclusif : celui du maniement du langage »[163]. Cela provoque chez lui une double sidération : sa fascination pour l'automatisme de l'objet, et le renvoi à son incomplétude, puisque cette machine est douée de facultés langagières qu'elle emploie dans un environnement virtuel dans lequel l'homme n'est qu'un visiteur, à l'instar de la dénomination employée pour désigner le trafic d'un site web. Comme l'exprime Jean Baudrillard dans son *Système des Objets*, « l'automatisme, c'est l'objet prenant une connotation d'absolu dans sa fonction particulière »[164]. Le philosophe prend l'exemple du levier de vitesse qui connote la vitesse absolument et complètement grâce à son automaticité. À cet idéal d'automaticité s'ajoute « une certaine marge d'indétermination, qui permet à la machine d'être sensible à une information extérieure »[165], que l'on retrouve dans le principe de *machine learning* (« apprentissage automatique ») qui, défini de manière simple, correspond à une « technologie d'Intelligence Artificielle permettant aux ordinateurs d'apprendre sans avoir été programmés explicitement à cet effet »[166]. Ce principe est au fondement d'Amazon Alexa, Google Assistant et SIRI, qui peuvent s'adapter aux requêtes et aux habitudes de consommation de leurs utilisateurs. Si nous n'avons pas les qualifications suffisantes

---

163 *INA Global*, Santolaria, N., « Interfaces vocales : attention, dangers ! », 06/12/2017. Disponible à l'adresse suivante : https://www.inaglobal.fr/numerique/article/interfaces-vocales-attention-dangers-10013

164 Baudrillard, J., *Le Système des Objets*, éd. Gallimard, 1968, p.153.

165 ibid., p.156.

166 *Le Big Data*, L., B., « Machine Learning et Big Data : Définition et Explications », 06/07/2018. Disponible à l'adresse suivante : https://www.lebigdata.fr/machine-learning-et-big-data

pour détailler les modalités du *machine learning*, un webinaire de Bloomberg[167], média et service financier américain, permet d'en comprendre les mécanismes plus précisément. En ce qu'elle fonctionne *a priori* seule, sans le moindre effort ou participation externe à l'élaboration de sa réponse, l'enceinte connectée fascine son utilisateur, qui y projette une conscience pourtant factice — une Intelligence Artificielle :

> « Parce que l'objet automatisé « marche tout seul », il impose une ressemblance avec l'individu humain autonome, et cette fascination l'emporte. Nous sommes devant un nouvel anthropomorphisme. Jadis les outils, les meubles, la maison elle-même portaient dans leur morphologie, dans leur usage, clairement empreintes la présence et l'image de l'homme. Cette collusion est détruite au niveau de l'objet technique perfectionné, mais il s'y substitue un symbolisme qui n'est plus celui des fonctions primaires, mais des fonctions super-structurelles : ce ne sont pas ses gestes, son énergie, ses besoins, l'image de son corps que l'homme projette dans les objets automatisés, c'est l'autonomie de sa conscience, son pouvoir de contrôle, son individualité propre, l'idée de sa personne »[168].

La construction d'une Intelligence Artificielle qui incarne la voix de l'enceinte connectée crée chez son utilisateur un sentiment de réciprocité mêlé de honte, puisque la machine présente une perfection de ses facultés : la voix intérieure de l'homme, support de sa conscience qui s'extériorise dans la vive voix, devient ici une voix mécanique qui projette l'intelligence de la machine. Ce sentiment de réciprocité est cependant déprécié par l'état de la technique, qui comme nous l'avons observé ne parvient aujourd'hui pas à créer une *persona* virtuelle capable de la même fluidité de parole et de compréhension que l'homme. Le progrès technique se compose ainsi de deux dimensions : l'état de son perfectionnement dans le moment présent, et un idéal de perfection qui se situe dans un futur inqualifiable. Ce dernier est comparable à « l'état positif » ou « scientifique » caractérisé par le philosophe Auguste Comte à l'époque de la révolution industrielle[169]. Selon lui, l'homme, qui ne peut connaître l'essence, la cause des phénomènes — le « pourquoi ? » — doit y substituer une appréhension du « comment ? » et savoir gouverner les phénomènes plutôt que d'en expliquer la cause. Cette descente se trouve concrétisée dans les enceintes connectées : l'homme applique à la machine les lois de la connaissance, de l'intelligence, de la conscience, qu'il sait maîtriser sans en connaître l'origine. Pour exemple, la voix conjugue deux qualités dont nous ne

---

167 *Bloomberg*, Rosenberg, D.S., « Foundations of Machine Learning ». Disponible à l'adresse suivante : https://bloomberg.github.io/foml/#home

168 Baudrillard, J., *Le Système des Objets*, éd. Gallimard, 1968, p.157.

169 Comte, A., *Discours sur l'Esprit Positif*, 1844, éd. Hachette Livre BNF, 2013.

sommes pas parvenus à expliquer l'origine métaphysique, et avons pour cette raison ramenée à un développement existentiel, à notre échelle humaine — celui du cri du nouveau-né qui s'ancre progressivement dans un langage articulé.

Dans ce contexte de déterminisme technologique à la fois idéologique et effectif, les enceintes connectées sont la preuve de la capacité de la science à créer artificiellement de l'intelligence. Les publics, attirés par sa promesse de simplification de leurs quotidiens, adoptent cette technologie pour ses qualités d'assistant. L'objectif est qu'à terme les publics intègrent cette technologie comme un prolongement naturel de leurs existences, que l'assistant devienne un compagnon. Pour ce faire, les concepteurs des Intelligences Artificielles doivent incarner leurs voix d'une personnalité, qui ne mobilise pas seulement une intelligence, mais une faculté d'émotion factice retranscrite dans la voix de la machine. L'alliance de l'intelligence et du sentiment est par conséquent un idéal de perfection vers lequel tendent les concepteurs, mais qu'ils ne parviennent à atteindre.

Cette problématique met en jeu une notion phare de la robotique : la *uncanny valley* (« vallée dérangeante »), théorisée par le roboticien japonais Masahiro Mori dans les années 1970. La « vallée dérangeante » consiste en ce moment où les technologies font montre de suffisamment d'automaticité pour se rapprocher de la figuration d'une conscience humaine, mais n'ont pas atteint la perfection qui permettrait leur adoption complète grâce à leur ressemblance avec l'homme — c'est la chute dans la vallée dérangeante, que le progrès technique s'efforce de franchir. La solution réside selon nous dans la capacité à donner une émotion aux machines, venant compléter l'intelligence et créer une personnalité artificielle. Une fois cet idéal atteint, un danger demeure pourtant : lorsque la machine imitera l'homme à la perfection, le moindre de ses défauts deviendra monstrueux, et suscitera le rejet d'un utilisateur devenu dupe. Des chercheurs arguent pour cette raison qu'avec le progrès technique, l'homme sera aussi de plus en plus capable de distinguer l'un de ses semblables d'une imitation robotique, rendant impossible de remonter tout à fait la vallée dérangeante — ce phénomène est qualifié de *uncanny wall* (« mur dérangeant »), traduisant ce caractère d'infranchissable[170].

La voix de l'Intelligence Artificielle dans les enceintes connectées consiste par conséquent en la recherche d'un équilibre entre la création d'une personnalité artificielle et la figuration de la

---

[170] *The Guardian*, Lay, S., « Uncanny valley: why we find human-like robots and dolls so creepy », 13/11/2015. Disponible à l'adresse suivante : https://www.theguardian.com/commentisfree/2015/nov/13/robots-human-uncanny-valley

machine. Nicolas Dot, chef de produit au pôle marketing du groupe radio RTL et chargé d'une veille sur les enceintes connectées, prend l'exemple du SIRI d'Apple :

> « Je pense que le côté artificiel ne va pas être le côté porteur des innovations. Le superficiel, l'artificiel a tendance à éloigner l'être humain de l'innovation. C'est le principe de la *uncanny valley* : à trop vouloir ressembler à l'être humain, à trop vouloir se rapprocher de sa réalité, on va s'éloigner de l'innovation. La voix SIRI, très hachée et robotique, suit cette logique. L'humain a besoin de se sentir maître de lui-même et de ne pas avoir du coup quelque chose en face de lui qui lui ressemble trop. Mais l'artificiel n'est pas un synonyme de défaite ou de risque. Ça permet à l'innovation de rester forte et de fédérer autour d'elle »[171].

Le caractère naturel de la voix des enceintes connectées doit par conséquent faire place à un juste dosage entre une fluidité d'expression permettant l'immersion de l'utilisateur, une capacité de compréhension médiation d'une interaction directe, et une sonorité à la fois naturelle et artificielle. Ce que Barthes appelle le « bruissement »[172] de la machine ne pourra se conjuguer d'un « grain » trace d'un référent sensible factice — une personnalité artificielle — que si elle parvient à faire preuve d'émotion. Cette capacité, pour l'instant absente des enceintes connectées, implique de leur créer une image de compagnons de vie par le biais d'un discours marketing dont l'axe est celui de la personnalisation des services rendus à l'utilisateur grâce à la supposée personnalité de la voix qui leur parle. Un exemple pour illustrer ce point : la publicité diffusée par Amazon pour ses enceintes Echo lors du Super Bowl en 2018 aux États-Unis[173]. Alexa, la personnalité artificielle de la firme de Jeff Bezos, y a perdu sa voix. De nombreuses célébrités tentent, de manière comique, de la remplacer sans succès. Alexa retrouve finalement sa voix, et prend le relais. Alexa est une personnalité, incarnée par son prénom et sa voix, qui imite l'homme et en dépasse les facultés. Elle est plus compétente que son utilisateur, justifiant son utilisation, et créant par là même une fascination pour son automaticité. Elle intègre les étapes du quotidien sans pour autant s'ancrer dans l'existence humaine car certains de ses traits rappellent son caractère artificiel.

Dans le contexte que nous avons dépeint, deux notions commandent la recherche technologique des concepteurs des Intelligences Artificielles intégrées aux enceintes connectées : la

---

171 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

172 Barthes, R., *Le Bruissement de la Langue*, éd. Seuil, 1984, p.99 : « Le bruissement, c'est le bruit de ce qui marche bien. Il s'ensuit ce paradoxe : le bruissement dénote un bruit limite, un bruit impossible, le bruit de ce qui, fonctionnant à la perfection, n'a pas de bruit ; bruire, c'est faire entendre l'évaporation même du bruit : le ténu, le brouillé, le frémissant sont reçus comme les signes d'une annulation sonore ».

173 YouTube, compte officiel Amazon, « Alexa Loses Her Voice - Amazon Super Bowl LII Commercial », 31/01/2018. Disponible à l'adresse suivante : https://youtu.be/J6-8DQALGt4

fluidité de leur expression et de leur compréhension, cruciales pour des technologies incapables de construire une personnalité complète, capable d'émotion.

b) Incorporer de l'humain dans la machine : l'expression et la compréhension, facteurs clés de la relation entre l'enceinte connectée et son utilisateur.

Deux éléments concentrent le travail des chercheurs employés par Amazon, Google, Apple et les autres groupes proposant des enceintes connectées : la fluidité de l'expression et la capacité de compréhension de leurs Intelligences Artificielles, qui constituent l'aboutissement des logiques d'immersion et d'interaction partagées par la voix et la virtualité numérique. Nous pourrions rassembler ces deux éléments dans la notion de conversation, qui comprend l'immersion dans une situation de communication immédiate et l'interaction qui la construit. La logique qui sous-tend ce travail de recherche et de conception est l'attribution de qualités humaines à la machine, pour l'immerger dans les existences de ses utilisateurs.

Lorsque l'on s'intéresse à ces procédés, un élément est frappant : la présence de l'humain à chaque étape de la communication de l'enceinte connectée. En effet, les solutions technologiques déployées pour permettre à la machine de communiquer supposent le travail conjugué d'ingénieurs et de dialoguistes afin d'organiser des subtilités d'expression uniquement perceptibles par l'homme dans des modèles systématiques. L'expression est ainsi l'apanage de profils littéraires, quand la compréhension est développée par des ingénieurs. Clothilde Chevet, doctorante du GRIPIC (Groupe de Recherches Interdisciplinaires sur les Processus d'Information et de Communication) CELSA Sorbonne Université, détaille dans un article publié sur le site *Méta-Média*[174] la composition des équipes de « plumes » qui œuvrent au sein de start-ups et de GAFAM pour que, au-delà du fait de parler le langage naturel, ils soient capables de « paraître humains ». Leur but, manier « une pré-écriture de l'oral et surtout du non-verbal, essentiel pour humaniser l'assistant et simuler un échange spontané », et « développer la corporalité de leur produit » grâce aux mots. En amont, des dialoguistes façonnent les réponses de l'assistant, qui seront déclenchées par une conversation déclenchée à la mention de leur *invocation name* (« Alexa… », « OK Google », « Dis SIRI » entre autres). Pendant la conversation, des *backers* qui veillent à l'harmonisation des paroles de l'assistant et évitent les défaillance. En aval, des « dresseurs d'AI *[Artificial Intelligence]* »

174 *Méta-Média*, Chevet, C., « Dialoguiste pour robot : l'humain derrière la machine », 06/06/2018. Disponible à l'adresse suivante : https://www.meta-media.fr/2018/06/06/dialoguiste-pour-robot-lhumain-derriere-la-machine.html

retranscrivent les échanges entre des utilisateurs et leurs enceintes connectées pour améliorer la qualité du service. Ces équipes fonctionnent au sein d'une boucle, où chacun enrichit et actualise le travail de l'autre.

Outre le vocabulaire employé, ces personnes œuvrent à la figuration d'un corps que la voix de l'Intelligence Artificielle viendrait extérioriser, à la manière de la vive voix, de deux autres manières. Premièrement, l'usage de détails semblables à ceux que Roland Barthes identifie dans le « grain de la voix »[175] que nous avons évoqué, parmi lesquels les onomatopées, dont les sonorités signifient l'ancrage dans un corps constitué d'émotions. Clothilde Chevet prend l'exemple du Google Assistant, dont une conversation téléphonique avec une coiffeuse a été pensée avec minutie :

> « La récente démonstration de Google lors de son congrès annuel le 8 mai 2018 a justement soulevé cette question en donnant à entendre un échange téléphonique entre l'Assistant Google et une employée d'un salon de coiffure. Durant cet échange, la femme au bout du fil a notamment demandé à l'assistant de patienter quelques secondes, ce à quoi l'assistant a répondu par un simple « Mmh mmh », déclenchant les rires du public assistant à la présentation. Cette onomatopée à elle seule révèle un pan entier de l'humanisation des machines »[176].

Deuxièmement, les dialoguistes œuvrent à représenter dans les réponses des assistants vocaux un corps — condition d'expression de la voix et de l'appréhension de la machine comme incarnant une personnalité. Celui-ci peut être figuré en référence à la machine, comme le Google Assistant — « Les déclarations d'amour touchent ainsi l'assistant « au plus profond de son code binaire », il a « l'algorithme dans la peau » et aime se balader le long des plages de Google Earth… Il ne peut donner de baiser, de peur de court-circuiter »[177] —, ou à l'inverse un corps humain, à l'instar de S Voice, l'ancien assistant vocal de Samsung — qui « allait jusqu'à proposer à l'utilisateur d'imaginer un contact physiquement impossible en disant « Ferme les yeux. Voilà, là, je te serre fort dans mes bras » lorsqu'on lui demandait un câlin. Il proposait également « d'embrasser (l'utilisateur) par la pensée » lorsqu'on lui demandait un baiser »[178].

---

175 Barthes, R., *Le Plaisir du Texte*, éd. Seuil, 1973, pp.104-105.

176 *Méta-Média*, Chevet, C., « Dialoguiste pour robot : l'humain derrière la machine », 06/06/2018. Disponible à l'adresse suivante : https://www.meta-media.fr/2018/06/06/dialoguiste-pour-robot-lhumain-derriere-la-machine.html

177 ibid.

178 ibid.

Plus largement, la fluidité de l'expression de l'assistant embarqué dans les enceintes connectées dépend de la capacité de ses concepteurs à lui inculquer une langue qui, attachée à un territoire, comprend des subtilités à la fois sonores et culturelles. Entre ici en jeu le deuxième aspect que nous souhaitons étudier : la compréhension, condition de la fluidité de l'expression de la machine, qui implique encore davantage les ingénieurs. Le lancement d'Amazon Echo en France a par exemple supposé un travail colossal de ses facultés de compréhension de la langue française, aux nuances distinctes de l'Anglais. Les accents régionaux, les artistes ou sports préférés d'Alexa, l'usage du pronom « vous », absent de l'Anglais, l'emploi fréquent du terme « week-end » aux accents francisés sont des exemples du nombre impressionnant d'éléments que les ingénieurs et les dialoguistes ont dû prendre en compte dans leur travail[179]. La conception de l'Intelligence Artificielle, régie par un idéal œcuménique de perfection de l'objet, est paradoxalement contrainte par les spécificités culturelles et linguistiques des territoires sur lesquels la technologie souhaite être implantée.

Parallèlement aux particularités territoriales, les concepteurs des Intelligences Artificielles qui animent les enceintes connectées doivent développer leur capacité de compréhension en tant qu'adaptation aux requêtes des utilisateurs. Cet aspect demeure le plus complexe, car il ne peut se limiter à un travail en amont — il se réalise dans le temps de l'interaction avec l'utilisateur. Si le *machine learning* (apprentissage automatique), que nous avons évoqué précédemment, semble être la solution à cet enjeu, ses avantages s'avèrent restreints. Ce type de mécanisme fonctionne efficacement lorsque les données que la machine doit engranger ne présentent aucune nuance — par exemple des données chiffrées, ou mathématiques. Or, la parole, qui dans la voix manie un référent sensible et un arbitraire du signe qu'elle contourne et combat dans le temps de son énonciation, présente trop de singularités pour que la méthode du *machine learning* soit pleinement efficiente. Le nombre d'évocations implicites qui se déploient dans les mots d'une conversation constitue un matériau trop complexe. Cette problématique va jusqu'à dépasser les compétences présentes au sein de Google, Microsoft ou Amazon. Ce dernier a par exemple créé un concours — le Alexa Prize[180] — dont la récompense de 3,5 millions de dollars sera remportée par l'équipe d'ingénieurs étant parvenue à concevoir une intelligence capable de tenir une conversation de vingt minutes sans la moindre anicroche. Les propositions avancées en 2018 nous ramènent au déterminisme

---

179 *Wired*, Barrett, B., « Inside Amazon's Painstaking Pursuit to Teach Alexa French », 13/06/2018. Disponible à l'adresse suivante : https://www.wired.com/story/how-amazon-taught-alexa-to-speak-french/?mbid=nl_061318_daily_list_p

180 La page de présentation du Amazon Alexa Prize est disponible à l'adresse suivante : https://developer.amazon.com/alexaprize

technologique caractérisant le domaine de l'Intelligence Artificielle (et par extension des appareils qui l'emploient), où le travail humain est encore nécessaire alors même que l'objectif est de l'automatiser, et de déplacer la voix d'un corps vers une machine. L'édition 2018 du concours a ainsi donné lieu à plusieurs détournements mêlant Intelligence Artificielle et travail humain, rapportés par un article de *The Verge*[181], site web américain spécialisé dans l'actualité des technologies et des médias. Pour exemple, l'équipe représentant le Royal Institute of Technology de Suède a mobilisé les services d'une plateforme appartenant à Amazon, Mechanical Turk, qui propose à des entreprises souhaitant développer leurs Intelligences Artificielles d'employer ponctuellement des humains pour des tâches explicitement décrites sur le site de la plateforme comme mieux réalisées par ces derniers, mais quelque peu répétitives et chronophages : *[« Alors que les technologies informatiques continuent de s'améliorer, il demeure de nombreuses choses que les êtres humains peuvent réaliser bien plus efficacement que des ordinateurs, comme identifier des objets dans une photo ou une vidéo, dé-dupliquer de la donnée, retranscrire des enregistrements audio ou chercher des détails dans des données. Traditionnellement, ces types de tâches sont accomplis en employant une large main d'œuvre à titre temporaire (ce qui prend du temps, coûte cher et est difficile à estimer) ou n'ont jamais été réalisées »]* :

> « While computing technology continues to improve, there are still many things that human beings can do much more effectively than computers, such as identifying objects in a photo or video, performing data de-duplication, transcribing audio recordings or researching data details. Traditionally, tasks like this have been accomplished by hiring a large temporary workforce (which is time consuming, expensive and difficult to scale) or have gone undone »[182].

Les ingénieurs de l'équipe, surnommée la « *Fantom Team* » pour l'occasion, ont employé les services d'humains par le biais de cette plateforme pour écrire des réponses adaptées à chacune des thématiques que l'utilisateur pourrait vouloir aborder lors d'une conversation avec l'assistant vocal. Face à ce constat, très courant dans le domaine, une journaliste du quotidien d'information britannique *The Guardian* a qualifié ce phénomène de la « montée de la pseudo-Intelligence Artificielle » (« *the rise of pseudo-AI* »), qui emploie des humains pour faire le travail de robots[183].

---

181 *The Verge*, Vincent, J., « Inside Amazon's $3.5 million competition to make Alexa chat like a human », 13/06/2018. Disponible à l'adresse suivante : https://www.theverge.com/2018/6/13/17453994/amazon-alexa-prize-2018-competition-conversational-ai-chatbots

182 Disponible à l'adresse suivante : https://www.mturk.com

183 *The Guardian*, Solon, O., « The rise of 'pseudo-AI': how tech firms quietly use humans to do bots' work », 06/07/2018. Disponible à l'adresse suivante : https://www.theguardian.com/technology/2018/jul/06/artificial-intelligence-ai-humans-bots-tech-companies

L'adoption complète des enceintes connectées par les publics fait par conséquent l'objet de plusieurs limites, que ses concepteurs entreprennent de dépasser : des difficultés d'expression et de compréhension qui tendent vers un idéal d'automatisme subissant la double pression d'une technologie trop imparfaite pour ressembler à l'être humain et qui emploie des hommes pour la faire fonctionner.

Notre analyse du marché des enceintes connectées sous des angles économiques, médiatiques et technologiques nous permet de conclure que la radio y occupe une place stratégique. Face aux autres marques médias, elle a l'avantage concurrentiel d'une identité ancrée dans le travail de la voix, affirmant sa légitimé sur ce support. Son matériau vocal constitue en outre l'aboutissement des processus d'immersion et d'interaction — lui conférant son effet d'authenticité — que le numérique emploie dans un espace virtuel. Face à l'encadrement de ses contenus par une voix artificielle, la radio peut faire ressortir l'authenticité de ses voix d'antenne et des corps qui, en coulisse, éditorialisent les programmes. La radio emploie des savoir-faire qui, s'ils altèrent eux-même les qualités naturelles du matériau vocal, créent un message nouveau incarné par un corps radiophonique. Sa place s'ancre dans un contexte où les concepteurs des Intelligences Artificielles emploient d'ailleurs bien plus le travail humain que nous pourrions le penser de prime abord. La radio, qui se déploie sur ce support dans une logique de média global, est cependant au centre d'une tension. Notre analyse de l'offre médiatique sur Amazon Echo en France laisse en effet percevoir un manque de créativité, qui peut autant provenir du faible développement de ce marché dans l'hexagone que d'une redéfinition des métiers de radio à l'ère numérique qui est encore en cours. Cela ne serait pas surprenant, car ce média est confronté à deux enjeux au sein de l'économie numérique, que les enceintes connectées illustrent et concrétisent. Premièrement, l'emploi de son avantage concurrentiel provenant des qualités immersives et interactives de son matériau vocal auquel elle confère une authenticité, source de son savoir-faire, sur le numérique peut avoir la conséquence contre-productive de diluer son identité de média ancré dans l'audio et le direct, au sein d'un écosystème de services emblème d'une stratégie de média global. Deuxièmement, la progressive artificialisation des modes de production et de consommation des contenus sur le numérique modifie la manière d'envisager et de réaliser les métiers de radio, dans un moment où ils constituent probablement l'atout principal de ce média. Ce sont ces deux aspects que nous souhaitons explorer tour à tour dans l'ultime partie de notre recherche.

III — Les enceintes connectées, indice d'une redéfinition du média radio entre menaces et opportunités.

Du paysage que nous avons dépeint transparaît un constat : la radio, en tant que média dont le travail de la voix est ancré dans son identité, est sans doute la plus légitime sur les enceintes connectées. Son innovation dans le domaine n'est pourtant pas la plus importante, comme notre analyse de l'offre des médias français sur Amazon Alexa semble le montrer. Cela peut néanmoins trouver une explication. La radio n'ayant pas besoin de créer des contenus audio pour ce support, elle préfère se reposer sur ses acquis, et par là même promouvoir son savoir-faire radiophonique, dans un contexte où le marché français des enceintes connectées demeure embryonnaire. La marque média se suffit pour l'instant à elle-même, et est juste déclinée sur les enceintes connectées. L'anticipation du succès de ce support pourrait être plus prégnante, à l'heure où les médias américains dédient des équipes à la quête de leurs voix propres, incarnations de leurs marques, dans un marché florissant. La radio doit trouver le juste dosage entre l'investissement dans une pluralité de supports de prise de parole et la figuration de ce qui fait sa spécificité — la voix et le son —, pour mettre en avant des marques médias qui émanent d'abord des métiers qui la composent. C'est dans cette optique que nous avons interrogé des personnes qui travaillent en radio, allant des travailleurs du son à des fonctions publicitaires et marketing, pour évaluer dans leurs professions la prégnance de cet enjeu. Si nous avons déjà cité plusieurs de nos entretiens au cours de notre analyse pour ancrer notre propos, nous nous focaliserons ici sur deux points : la construction d'une vision du monde axée autour de la voix et du son comme fondation des métiers de radio, et la redéfinition de ce que Hervé Glévarec, directeur de recherche au CNRS, qualifie de « pragmatique radiophonique »[184], c'est-à-dire la manière dont cette vision du monde est employée pour créer les contenus de radio. Pour ce deuxième point, notre angle d'approche sera celui d'une potentielle redéfinition des métiers de radio face à l'émergence des enceintes connectées, à la fois transistor augmenté d'une possibilité d'interaction et indice d'un mouvement d'artificialisation des professions médiatiques.

L'ultime partie de notre étude comprendra trois mouvements : la place des voix de radio au sein d'une économie numérique que nous avons entrepris de décrire plus tôt en se focalisant sur les

---

184 Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001, p.15.

enceintes connectées, la perception que les personnes travaillant en radio ont de l'évolution de leurs métiers dans ce contexte, puis la formulation de recommandations à destination des stations de radio. L'analyse du statut de la radio dans l'environnement que nous avons caractérisé lors de nos précédentes parties devrait en effet nous permettre d'établir des recommandations stratégiques pour les stations de radio, qui sont au carrefour des enjeux mis en lumière par l'émergence des enceintes connectées à l'échelle internationale.

A. LA PLACE DES VOIX DE RADIO AU SEIN DE L'ÉCONOMIE NUMÉRIQUE : ENTRE AVANTAGE CONCURRENTIEL ET PERTE D'IDENTITÉ.

1. L'économie de la voix numérique : des propriétés d'immersion et d'interaction intégrées à un champ concurrentiel élargi par la possibilité d'une incarnation.

Sur le numérique, la voix de radio circule au sein de trois mouvements : la virtualité, avec laquelle elle partage ses qualités d'immersion et d'interaction immédiate ; un espace dominé par l'image, qui suppose d'y ajouter de la vidéo ou des audiogrammes ; son utilisation par une multiplicité d'acteurs, qui a la conséquence de la voir se déplacer d'un savoir-faire radiophonique vers d'autres médias et vers des Intelligences Artificielles. C'est sur ce dernier point que nous souhaitons nous attarder. Si nous en avons étudié les modalités — l'altération de la source de la voix et la construction d'un champ concurrentiel élargi —, nous observons que les propriétés de notre matériau se voient attribuer une mission qui dépasse l'enjeu de diversification des audiences d'une pluralité de marques. En effet, sur le numérique, la voix devient une possibilité de donner corps à une entité médiatique, de l'incarner — de la même manière que la vive voix d'un individu projette son intériorité, sa personnalité dans le monde. De nombreuses marques — commerciales ou médias — œuvrent ainsi à la construction d'une voix unique, qui les représente et leur confère une substance dans le monde virtuel, dans le double objectif de se différencier et de fidéliser des audiences. Cette quête d'incarnation prend des formes multiples, et est mise en œuvre sur l'ensemble des supports audio numériques (les enceintes connectées et les podcasts étant les principaux). Le média peut faire le choix de créer une voix artificielle qui lui est propre, à la manière de Huffie du *HuffPost*, ou d'engager des personnes réelles pour incarner sa marque, comme la radio qui a ses voix d'antenne. Cela suppose à chaque fois de réorganiser les métiers en interne,

afin de produire des contenus dédiés à l'audio et aux usages qu'ils proposent, qui entrent plus ou moins en résonance avec les supports d'origine de ces médias.

Les États-Unis, lieu de la renaissance du podcast (initiée en 2014 avec *Serial*)[185] et berceau des enceintes connectées, constituent le marché où cet enjeu est le plus vif. Nombre de médias du pays — Al Jazeera, groupe qatari, faisant office d'exception dans les exemples que nous allons citer, bien qu'il soit implanté sur ce territoire — créent des équipes dédiées à la création de formats audio innovants pour les enceintes connectées et à la quête d'une voix unique, signature de chaque marque sur le numérique. Un article de *Digiday*[186], *pure player* spécialisé dans l'actualité média et marketing, rapporte plusieurs exemples pour l'année 2018. NPR (*National Public Radio*, radio publique nationale des États-Unis) a une équipe de six personnes qui se consacrent aux assistants vocaux. Le *New York Times* est à la recherche d'un « *Voice Editor* », sorte de rédacteur en chef pour les technologies vocales. Al Jazeera a engagé une équipe travaillant à réutiliser des contenus existants sur les enceintes connectées et à en créer de nouveaux. Au *Washington Post*, six personnes travaillent à plein temps sur le développement de formats audio. Bloomberg œuvre quant à lui à la conception de contenus dédiés aux assistants vocaux. La voix fait l'objet d'un fort intérêt de médias dont l'écriture constitue encore le cœur de métier, et dépasse les enceintes connectées. Pour exemple, le *New York Times*, qui compose une équipe dédiée aux assistants vocaux, fait déjà de nombreuses incursions dans le podcast, allant du documentaire (notamment « *Caliphate* », sur les décombres de l'État Islamique après de prise de Mossoul, en Irak) à l'actualité quotidienne (avec « *The Daily* », programme d'une vingtaine de minutes) en passant par les thématiques de société — par exemple l'amour dans « *Modern Love* »[187]. Des radios comme la NPR se saisissent quant à elles de cette opportunité pour construire une voix numérique concrétisée dans des contenus courts dédiés aux enceintes connectées, sans être certaines de générer des audiences suffisantes à la monétisation de leurs contenus.

Outre les enceintes connectées et les podcasts, la voix intègre les pratiques de médias écrits, qui font à nouveau l'objet d'une dépendance aux concepteurs des technologies vocales et des plateformes qui organisent le contenu, souvent détenues par des GAFAM. Dans ce cadre, la voix est soit artificielle, soit physique, mais répond toujours à l'objectif d'insuffler un ton singulier à un

---

185 *New York Magazine - Daily Intelligencer*, « What's Behind the Great Podcast Renaissance? », 31/10/2014. Disponible à l'adresse suivante : http://nymag.com/daily/intelligencer/2014/10/whats-behind-the-great-podcast-renaissance.html

186 *Digiday*, Moses, L., « For news publishers, smart speakers are the hot new platform », 23/08/2018. Disponible à l'adresse suivante : https://digiday.com/media/news-publishers-smart-speakers-hot-new-platform/

187 Les descriptions de ces programmes sont disponibles à l'adresse suivante : https://www.nytimes.com/podcasts/

média global qui s'exprime sur une multitude de supports. La lecture d'articles par des voix artificielles fait par exemple partie prenante de la stratégie de Bloomberg, groupe spécialisé dans les services et l'information économique. Le média utilise une solution logicielle de *text-to-speech* développée par Amazon — Amazon Polly — pour rendre accessible tous ses articles écrits en audio[188]. Symptôme d'une convergence des pratiques amateures et professionnelles dans la production médiatique identifiée par Henry Jenkins[189], et encadrée par les GAFAM, Amazon a rendu son logiciel Polly accessible à tous les éditeurs d'un blog ou d'un site développé avec WordPress sous la forme d'un *plug-in*[190], en février 2018[191]. En guise de contrepied, des médias privilégient la lecture de leurs articles par des comédiens. Les « *spoken editions* » (« éditions parlées »), diffusées sur les plateformes de podcast et de *streaming* musical (qui croissent en intégrant un nombre maximal de fonctionnalités et de contenus), ont d'ailleurs créé une économie des voix, qui emploie des comédiens dédiés à la lecture numérique[192]. Cette économie met en concurrence directe des start-ups et des groupes valorisés à plusieurs milliards de dollars. Pour exemple, Audm, une plateforme américaine proposant à des médias d'engager des comédiens pour lire leurs formats longs (*long reads*) à la manière d'un récit, incarné, les distribue sous la forme d'un service payant (6,99 dollars par mois). Elle fait l'objet de la concurrence de GAFAM tels qu'Amazon, qui étend son service de livres audio — Audible — à des contenus issus de la presse[193], ou qu'Apple, avec son service SpokenEdition[194]. Notons que l'usage de l'oralité n'est pas réservé aux entreprises médiatiques. Dans le même mouvement de convergence que nous avons évoqué,

---

188 *Ted Crunch*, Ha, A., « Bloomberg Media Group's Chief Product Officer sees big opportunities in audio », 07/08/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/06/07/bloomberg-doubles-down-on-audio/?guccounter=1

189 Jenkins, H., *Convergence Culture: Where Old and New Media Collide*, New York University Press, 2006.

190 La page de présentation du plug-in Amazon Polly est disponible à l'adresse suivante : https://wordpress.org/plugins/amazon-polly/

191 *Ted Crunch*, Perez, S., « Amazon launches a Polly WordPress plugin that turns blog posts into audio, including podcasts », 08/02/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/02/08/amazon-launches-a-polly-wordpress-plugin-that-turns-blog-posts-into-audio-including-podcasts/

192 *Columbia Journalism Review*, Brey, J., « Amid audio boom, magazines hire actors to voice their work », 05/01/2018. Disponible à l'adresse suivante : https://www.cjr.org/business_of_news/actors-magazines-voice.php

193 *The New York Times*, Chokshi, N., « Amazon's Audible Goes Long on Short-Form Audio », 07/07/2016. Disponible à l'adresse suivante : https://www.nytimes.com/2016/07/08/business/media/amazons-audible-goes-long-on-short-term-audio.html

194 *Ted Crunch*, Perez, S., « Spoken Editions goes live on iTunes so you can listen to your favorite websites », 03/10/2016. Disponible à l'adresse suivante : https://techcrunch.com/2016/10/03/spoken-editions-go-live-on-itunes-so-you-can-listen-to-your-favorite-websites/

Facebook teste par exemple en Inde une fonctionnalité permettant aux utilisateurs de publier de courts contenus parlés, nommée Voice Clips[195].

La voix est par conséquent un objet de convoitise pour l'ensemble des médias, que le cas des enceintes connectées illustre et concentre. Pour des médias dont les identités sont ancrées dans l'écriture, la voix est synonyme de nouvelles consommations, et permet de déployer la présence de leurs médias globaux. Le matériau vocal conjugue l'incarnation d'un discours dans un corps et la possibilité d'écoutes en situation d'accompagnement, qui est l'une des caractéristiques historiques de la radio que nous avons identifiées.

2. La remédiation du média radio dans l'économie numérique de la voix : la place des stations historiques au sein de l'audio numérique.

L'économie numérique de la voix rassemble des intérêts, des positions et des maniements multiples de ce matériau, dont les enceintes connectées constituent selon nous un aboutissement. Dans ce contexte, les fonctions et les effets de la radio ne lui sont plus propres. Le média radiophonique n'est plus l'apanage des stations historiques. Ce mouvement, que nous avons tenté d'analyser, constitue sa remédiation, définie comme le processus où des médias nouveaux pour leur époque « importent d'anciens médias dans l'espace digital pour en faire la critique et les remodeler »[196] *[« (...) they import earlier media into a digital space in order to critique and refashion them »]*. Si la radio en tant que substance — dont la nature est le transport du matériau vocal dans un espace médiatique — persiste, les stations s'affaiblissent, et leurs modes de fonctionnement font l'objet de l'appropriation d'autres acteurs, qui les altèrent (par exemple en proposant des contenus délinéarisés, ce que ne faisait pas la radio). Un article de *Syntone*, site web dédié à la couverture de l'art radiophonique, fait ainsi le constat d'une « évolution actuelle de la radio vers l'audio [qui] laisse notre flux hertzien traditionnel à l'arrière-plan, comme un vieux monsieur dépassé »[197]. Nous allons tenter d'observer en quoi cette remédiation, effective à l'échelle du média radio (a), suppose pour chaque marque qui le compose de travailler une image spécifique mobilisant le travail de la voix et des identités distinctes en termes de positionnement éditorial (b), mais également d'utiliser les technologies d'Intelligence Artificielle non pas en réaction mais au

---

195 *Ted Crunch*, Constine, J., « Micropodcasting? Facebook tries Voice Clip status updates », 02/03/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/03/02/facebook-voice-clips/

196 Bolter, J.D., Grusin, R., *Remediation. Understanding New Media*, MIT Press, 1998, p.52.

197 *Syntone*, Noiseau, E., « Vers un art radiophonique numérique ? », 18/04/2013. Disponible à l'adresse suivante : http://syntone.fr/vers-un-art-radiophonique-numerique-4-4/

service des métiers composant le savoir-faire radiophonique, bien que cet idéal soit contraint par les hiérarchies à l'œuvre dans les marchés numériques (c).

a) La perte de monopole du média radio sur le numérique.

La remédiation de la radio en tant que substance médiatique menace plusieurs spécificités des stations qui l'incarnent historiquement, dont la diffusion en direct ne représente qu'un aspect. Hervé Glevarec, que nous avons mentionné en ouverture de cette partie, consacre un article au « nouveau statut de la radio dans l'environnement numérique » dans la revue *Le Temps des Médias*[198]. Il identifie quatre fonctions historiques propres au média radio et deux menaces qui le guettent. Ces quatre fonctions constituent pour la radio des spécificités que des plateformes et des médias remédient dans l'économie numérique, et qui fondent les effets que nous avons réunis sous le terme d'authenticité. Selon Hervé Glévarec, les radios généralistes et thématiques remplissent tout d'abord une « fonction de sphère publique et d'identification » pour leurs auditeurs, qui en tant que consommateurs d'un média appartiennent à une communauté sociale et politique réunie par ses affinités, sans pour autant dénaturer « la visée d'intérêt général de cet espace ». Une « fonction d'actualité » est elle aussi au cœur du média radio, qui transmet l'information de manière immédiate. La radio est par ailleurs le média d'une « co-présence » pour un auditeur participant et s'inscrivant dans un public. Elle construit un lien social présent dans d'autres médias tels que la télévision, dont la dimension « cérémonielle »[199] est accentuée lors des événements importants. La radio remplit finalement une « fonction de programmation » correspondant à l'édition de ses contenus et l'établissement d'une offre spécifique.

En nous référant aux effets d'intimité, de communauté, de réel et d'imaginaire (ou d'évocation) de la radio que nous avons rassemblés sous la notion d'authenticité, ces quatre fonctions en constituent l'organisation au sein du dispositif radiophonique, mise en ondes du matériau vocal. Les fonctions « de sphère publique et d'identification », d' « actualité » et de « co-présence » caractérisent une parole de radio qui s'ancre dans le présent de son énonciation, historiquement diffusée en direct, s'adressant à la fois à un auditeur (effet d'intimité) et à un public (effet de communauté) avec la voix comme seul matériau (effet d'imaginaire amplifié par un dispositif acousmatique) constitutive d'un récit radiophonique à l'effet de réel (car la voix est la

---

198 Glévarec, H., « Le Propre de la Radio. Fonctions Radiophoniques et Nouveau Statut de la Radio dans l'Environnement Numérique », in *Le Temps des Médias*, n°22, 2014.

199 Dayan, D., Katz, E., *La Télévision Cérémonielle. Anthropologie et Histoire en Direct*, éd. PUF, 1996.

seule source du récit, et son grain lui donne une matérialité). La « fonction de programmation » consiste quant à elle en l'organisation du flux radiophonique en une grille et des voix pour l'incarner.

Hervé Glévarec distingue deux menaces pour les stations de radio, dans un mouvement de remédiation de leur substance en contexte numérique, où d'autres médias s'accaparent son matériau vocal, allant jusqu'à en déplacer la source dans le cas des Intelligences Artificielles qui animent les enceintes connectées. La fonction de « co-présence » est pour le chercheur remédiée dans les réseaux sociaux. Nous avons en effet constaté que si la radio et ces réseaux fonctionnent dans une relative harmonie, c'est parce qu'ils partagent leurs propriétés d'instantanéité et d'interactivité, qui fondent des communautés mais menacent la prégnance de la radio dans les usages des publics. La délinéarisation des programmes, que nous avons vue à l'œuvre dans les podcasts et les contenus courts dédiés aux enceintes connectées, constitue pour Hervé Glévarec l'ultime menace à laquelle fait face la radio. Cette dernière, dont l'identité se définit historiquement par la notion de direct, qui lui confère les fonctions précédemment détaillées, fait face à la concurrence des podcasts natifs (créés en France par des studios indépendants tels que Binge Audio, Louie Média ou Boxsons), qui l'amène à redéfinir sa manière de s'exprimer. Les stations de radio s'emparent ainsi du podcast, altérant leurs fonctions de deux manières — une première identifiée par Hervé Glévarec, et une seconde que nous ajoutons. Premièrement, lorsqu'elle propose un programme de sa grille en rattrapage, la radio dénature son flux en direct, risquant de la reléguer à une fonction de « radiothèque ». Les podcasts, « parfois aux frontières du comique de situation » pour le chercheur, consistent en l'action paradoxale de découper un flux en direct en morceaux distribués dans des temporalités décorrélées de celles de leur conception. Les émissions de France Culture, par exemple, sont parfois enregistrées en une matinée puis divisées en plusieurs épisodes distribués chaque semaine. La grille de programmes et sa diffusion en direct sont remises en question, créant « davantage une radio à la carte qu'une radio de rendez-vous ». Deuxièmement, la radio conçoit des programmes uniquement pour le numérique en réaction aux offres des studios de podcasts natifs, leur niant une place dans son flux historique. Pour prolonger l'exemple de France Culture, la station a créé le 13 février 2018 la série de fiction « Hasta Dente ! »[200], dédiée à l'écoute numérique.

Si les réseaux sociaux et la délinéarisation sont des menaces pour le fonctionnement historique de la radio, nous remarquons que les stations traditionnelles s'en emparent à raison. En effet, les stations de radio s'insèrent dans une économie numérique régie par la notion de média

[200] La page de présentation du podcast natif de France Culture « Hasta Dente ! » est disponible à l'adresse suivante : https://www.franceculture.fr/evenement/hasta-dente

global, qu'elles ne peuvent nier. La place de la radio se caractérise ainsi par une tension : la quête d'un juste équilibre entre l'adaptation aux usages numériques et le travail d'une image de marque dont la valeur ajoutée réside en sa spécificité — le travail de la voix en tant que composante du son. À la remédiation de la substance radiophonique se substitue par conséquent le travail d'une image de marque pour chaque station de radio, dans un environnement concurrentiel élargi qui altère son matériau vocal.

b) Un travail de l'image de marque entre mobilisation des savoir-faire radiophoniques et la représentation d'identités spécifiques.

Si les fonctions du média radio identifiées par Hervé Glévarec sont au fondement de sa spécificité au sein du champ concurrentiel élargi du numérique, elles l'ancrent également dans un rôle social. Ce dernier est mis en lumière par la notion de média de masse, dont les audiences se fragmentent sur le numérique. La radio est un instrument de culture au carrefour duquel circulent des informations et des idéologies, qu'elle diffuse à des communautés d'auditeurs. Si le média radio œuvre à sa rémunération grâce à des espaces dédiés aux messages d'annonceurs, sa fonction sociale diffère de celle de ces derniers. La séparation plus ou moins hermétique entre une régie publicitaire et les programmes d'un média en atteste, surtout lorsqu'il s'agit d'information. Une indépendance des rédactions de journalistes est conservée, même si elle est parfois mise en doute, supposant que ses membres s'en justifient au risque de rompre le contrat qui les unit à leurs audiences. Un média, en ce qu'il se compose d'un public auquel il présente une offre avec laquelle il se rémunère, a les attributs d'une marque commerciale, mais il ne place pas cette rentabilité au premier plan de ses préoccupations. En outre, un média se compose d'une pluralité de voix, de signatures et de tonalités éditoriales, une polyphonie dont l'antenne d'une radio constitue un exemple. Il se distingue d'une marque commerciale qui n'a qu'un seul référent — une stratégie de communication articulée autour d'un logo et d'un slogan, dont les campagnes publicitaires ne constituent qu'une déclinaison et non une prise de parole en propre. Ce n'est pas une transaction commerciale mais un contrat de communication porté par un dispositif — dont nous avons précédemment défini les modalités — qui unit un auditeur à une situation d'écoute radiophonique.

Si la notion de marque en tant que vecteur de distinction comprenant « une valeur, un savoir-faire, une expertise, une histoire, un engagement, une caution »[201] permet de construire un

[201] Glossaire E-Marketing, définition de « marque ». Disponible à l'adresse suivante : http://www.e-marketing.fr/Definitions-Glossaire/Marque-238333.htm

socle à partir duquel un média communique sur lui-même dans le but de se valoriser auprès des publics et des annonceurs, elle n'est par conséquent pas applicable en tous points aux médias. Sa pleine et entière mobilisation pour une entreprise médiatique lui fait courir le risque d'un « aplatissement de la polyphonie inhérente à tout média, [de l']affadissement de l'image par la sélection des traits caractéristiques les plus en phase avec les tendances socioculturelles censées être les plus appréciées par les publics et par les annonceurs, [et d'une] mise en situation d'équivalence avec des marques-enseignes et des marques-produits qui n'ont pas le même rôle social bien qu'elles y prétendent de plus en plus souvent »[202].

Forts de ce constat, nous pourrions arguer que la manière dont la radio peut se distinguer au sein d'un univers concurrentiel numérique qui s'accapare son matériau vocal consisterait en la stricte articulation des termes composant l'expression « marque-média ». Le travail d'une image de marque, à l'échelle d'une station, pour véhiculer sa spécificité éditoriale ; un travail de représentation du média d'origine, la radio, pour distinguer le travail radiophonique de celui concomitant des autres médias globaux et des Intelligences Artificielles en contexte de virtualité numérique. Si les modalités de la construction de l'identité d'un média radiophonique sont un premier élément de réponse à la problématique de distinction à laquelle les stations de radio font face, elles ne peuvent suffire. En effet, la radio — comme nombre d'autres médias — subit les évolutions des usages sur le numérique, car elle n'en est pas à l'origine. La radio filmée et la voix artificielle des enceintes connectées sont des illustrations de cette contrainte exercée par les GAFAM sur notre média, qui évolue non pas en anticipation mais en réaction à ces changements. La radio, plus que de s'y distinguer, doit mobiliser les propriétés de la virtualité numérique pour s'affirmer comme force de proposition. Ce mouvement s'effectuerait dans un contexte favorable, puisque la voix et la virtualité partagent des qualités d'immersion et d'interaction.

c) La mobilisation des technologies de l'espace virtuel pour nourrir les savoir-faire qui constituent les marques de médias radios.

Cette mobilisation, la radio l'effectue en partie. Sa réactivité et sa capacité à fédérer des communautés, qu'elle partage avec les réseaux sociaux, rend sa communication efficace sur ce support. Sur les enceintes connectées, notre analyse de l'offre des stations de radio pour Amazon Alexa en France semble montrer qu'elles ne se sont pas appropriées cette possibilité d'interaction

[202] Patrin-Leclère, V., « Un média est-il une marque ? », in *Communication*, Vol. 32/2, 2013, mis en ligne le 07 avril 2014. Disponible à l'adresse suivante : https://journals.openedition.org/communication/5027

directe avec l'usager dans une situation d'immersion vocale et sonore. La duplication des contenus proposés à l'antenne a d'avantageux qu'elle représente un savoir-faire radiophonique — une sorte de produit d'appel —, mais de nuisible qu'elle limite l'homogénéité de son expression sur ce support. Sa possibilité d'adaptation aux usages et aux configurations des enceintes connectées n'est par ailleurs qu'éphémère. Les médias s'y expriment aujourd'hui en situation d'horizontalité, sans que l'un prenne le pas sur l'autre. Un marché, composé de *leaders* et de prises de paroles alternatives, finira probablement par s'organiser. Les enceintes connectées, en limitant l'interaction avec les contenus à la voix, ont de singulier qu'elles réduisent les possibilités de réponses. Un article de *Quartz*, média spécialisé dans l'actualité économique, qualifie ce phénomène de « loyauté accidentelle » (« *incidental loyalty* »)[203] : l'enceinte connectée, qui facilite l'acte de consommation en ne mobilisant que la voix, crée des habitudes chez les utilisateurs provenant des suggestions qui leur sont faites. Ces suggestions sont issues du bon référencement d'une marque sur Internet, de partenariats noués avec les plateformes d'Amazon, Google ou Apple (comme franceinfo sur les enceintes Echo en France), ou d'un réflexe créé par un média ou une marque vers son service, lié à la construction de son identité.

Les technologies d'Intelligence Artificielle, si elles présentent de nouveaux types de relations à nouer avec les publics, sont constituées d'algorithmes facilitant la personnalisation du contenu. En travaillant leurs mécanismes pour aider les utilisateurs à composer des offres qui leur ressemblent et les journalistes à créer des contenus au haut degré d'affinité avec leurs publics, elles pourraient affirmer la place des entreprises médiatiques sur le numérique comme forces de proposition[204]. Cet idéal fait face à l'accaparement des données par les concepteurs de ces algorithmes, et au risque d'un enfermement de la curiosité de l'utilisateur dans une offre médiatique constituée grâce des traits saillants de ses goûts identifiés mécaniquement, en cela standardisés. C'est pourtant dans ce contexte sans issue apparente que la radio a selon nous une carte maîtresse à jouer : mobiliser ce qui fait sa force, c'est-à-dire les personnes qui travaillent le matériau vocal tant dans ses sonorités que dans la conception de propositions éditoriales, pour signifier son incarnation humaine. Si la virtualité numérique partage les propriétés d'immersion et d'interaction de la voix, elle en accélère la portée en même temps qu'elle en abaisse l'intensité. La chaleur, le corps que le grain de la voix de radio parvient à transmettre, appuyé par une légitimité dans son maniement

---

203 *Quartz*, Gerhart, A., « Alexa makes decision-making easier than ever—by making your choices for you », 10/07/2018. Disponible à l'adresse suivante : https://qz.com/1323884/alexa-makes-decision-making-easier-than-ever-by-making-your-choices-for-you/

204 *Ted Crunch*, Koponen, J., « A new hope: AI for news media », 12/07/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/07/12/a-new-hope-ai-for-news-media/

provenant de son histoire, est finalement l'atout majeur de ce média. Plus que l'enjeu d'un déploiement d'un média global sur les enceintes connectées qui subit l'encadrement par une voix artificielle ou en joue, c'est la perception que les personnes travaillant en radio ont de leurs métiers et de ses évolutions qui déterminera l'axe stratégique du paysage des stations de radio françaises sur le numérique.

B. LA PRAGMATIQUE RADIOPHONIQUE À L'ÈRE NUMÉRIQUE : LES ENCEINTES CONNECTÉES, INDICE D'UNE REDÉFINITION DES MÉTIERS DE RADIO.

En analysant la place de la radio sur les enceintes connectées, nous avons saisi l'ampleur de sa redéfinition au sein d'une économie numérique emplie de paradoxes. Spécificité issue d'un support d'origine ou média global ? Voix naturelle ou artificielle ? Corps physique ou corps virtuel ? Marque ou média ? La réponse est chaque fois ambivalente, et concentrée dans les qualités du matériau vocal, lui-même théâtre d'une tension entre musicalité des émotions et arbitraire du signe linguistique. La place de la radio sur les enceintes connectées dépend ainsi de la perception de ces enjeux par les personnes qui la créent, et travaillent la dématérialisation du grain de la voix au quotidien. Nous nous sommes pour cette raison entretenus avec neuf personnes aux profils divers, toutes employées du groupe radio RTL, qui a été absorbé par M6 au sein d'un actionnariat global de l'Allemand Bertelsmann en 2017[205].

Ces entretiens semi-directifs rassemblent des personnes aux fonctions variées. Si leurs apports respectifs ont été détaillés dans la partie de l'introduction à notre recherche dédiée à la présentation de notre démarche méthodologique, il nous semble important de les rappeler ici. Julien Le Ny et Alban Rivoal, producteurs à Fun Radio, ainsi que Rafaële Dhuicque-Filloux, fondatrice de la société de production de messages publicitaires radio Zéro Décibel, travaillent directement la matière sonore en écrivant des textes, en interagissant avec des comédiens, en menant des séances d'enregistrement, et en participant à leur montage. Les fonctions des régisseurs-monteurs de la station RTL délégués aux Opérations Publicitaires Spéciales (OPS), Vladimir Bobrow et Sacha Ephraim, se concentrent sur ces deux derniers points. Nous avons adjoint aux propos de ces cinq

[205] *Le Monde*, « M6 va racheter les radios de RTL pour 216 millions d'euros », 14/12/2016. Disponible à l'adresse suivante : https://www.lemonde.fr/actualite-medias/article/2016/12/13/m6-ouvre-des-negociations-pour-racheter-rtl-group-en-france-pour-216-millions-d-euros_5048419_3236.html

travailleurs du son ceux de personnes travaillant aux pôles OPS et marketing du département radio de la régie M6 Publicité, qui ont été interrogés pour leur vue stratégique du secteur. Guillaume Duhamel, concepteur-rédacteur d'opérations spéciales et Jérémy Valet, co-directeur de M6 Unlimited, structure dédiée aux opérations spéciales *crossmedia* (mobilisant plusieurs actifs du groupe) ajoutent à leurs oreilles d'auditeurs avertis une vision commerciale du secteur de la radio, dont leur travail consiste à contribuer à la rentabilité avec des dispositifs en forte affinité avec les programmes. Eux aussi interrogés pour leur connaissance du média radio, les chefs de produit marketing auxquels nous nous sommes adressés intègrent la voix et le son au cœur de leurs fonctions. Pierre Laurent, à la tête du Laboratoire Sonore, utilise des données relatives à la perception du son à des fins d'efficacité publicitaire. Nicolas Dot est chargé d'une veille sur les enceintes connectées, symptôme de la perspective de redéfinition du média radio que nous avons entrepris de dépeindre.

Nous souhaitons avec ces entretiens évaluer la prégnance de deux enjeux, que nous avons analysés d'un point de vue extérieur dans les parties précédentes. La radio fait selon nous l'objet d'une redéfinition, que les enceintes connectées accélèrent car elles déplacent la source du matériau vocal d'un corps humain à une Intelligence Artificielle. C'est sous le prisme de la voix, fil conducteur de notre recherche, que nous avons défini les deux aspects de notre étude de ces entretiens :

1) **La prégnance de la voix comme élément de définition d'une « pragmatique radiophonique »**[206], c'est-à-dire une vision du monde articulée autour du son, et ayant pour mécanismes principaux les notions de forme et de mise en forme présentes dans le processus de mise en ondes.

2) **La perception des enceintes connectées comme un objet de redéfinition des métiers de radio**, car directement axé sur la voix dans un environnement numérique où ce matériau est accaparé par une multitude d'acteurs et est déplacé vers des technologies, symptôme d'une progressive artificialisation des pratiques médiatiques.

L'étude de ces deux aspects, que nous allons aborder tour à tour, a pour objectif de saisir la perception que les personnes travaillant en radio ont de l'évolution de leur média, et le degré

[206] Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001, p.15.

d'importance qu'ils accordent aux enceintes connectées, transistor augmenté de la possibilité d'une interaction.

1. La prégnance de la voix comme élément de définition d'une « pragmatique radiophonique ».

Mêlant de l'observation participante, des entretiens et des données statistiques, Hervé Glévarec, directeur de recherche au CNRS, a publié en 2001 une étude sociologique des métiers de l'antenne de France Culture[207]. Dans cet ouvrage qui dévoile les coulisses des professions de travailleurs du son (parmi lesquels le monteur, le réalisateur, ou le producteur radio), le chercheur instaure plusieurs notions que nous souhaitons présenter, avant de nous les approprier pour analyser nos entretiens menés avec des membres du groupe radio RTL. Selon lui, les notions de forme (le matériau vocal) et de mise en forme (faisant écho à nos considérations sur le « dynéma » de Pierre Schaeffer dans une perpective de mise en ondes de la voix et du son) sont fondatrices d'une conception des émissions de radio comme des « actions », qui consistent en la « mise en œuvre de dispositifs » radiophoniques[208]. Au sein de cette « activité professionnelle mettant en présence, pour un temps donné, différentes fonctions et acteurs »[209] se construit une « socialisation professionnelle des personnels de France Culture [qui] consiste ainsi dans la mise en place d'une « vision du monde » nouvelle qui s'articule autour du son »[210]. C'est grâce à cette vision du monde que les professionnels de radio déploient une « structuration de leur objet » — l'émission en tant que matériau brut composé de voix — qu'ils peuvent opérer en équipe car ils en partagent la même perception : « « émission déséquilibrée », si des séquences de trois minutes font suite à vingt minutes d'entretien ; « ça tombe », en parlant de ce qu'on ne retient pas dans une bande enregistrée ; « on peut sauver ça », à propos de quelques minutes jugées intéressantes pour l'émission ; « type bon, type pas bon », pour parler de la qualité mixte, sonore et intellectuelle, d'un interviewé »[211]. Ces catégorisations du matériau sonore, dont la voix constitue l'une des composantes principales, construisent un « monde radiophonique » qui, converti en l'action de mise en œuvre d'un programme, devient une « pragmatique radiophonique » qui « se caractérise par des formes

---

207 Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001.

208 ibid., pp.10-11.

209 ibid., p.11.

210 ibid., p.14.

211 ibid., p.15.

d'engagement des locuteurs, de structuration de l'émission, d'attention à la langue et au son, particulières »[212]. La voix de radio est pour les personnes qui œuvrent à la construction de ce média un matériau qui fonde la pratique de leurs métiers, tant dans son maniement sonore (comme le « dynéma » de Pierre Schaeffer entreprend de le théoriser) que dans son insertion dans un dispositif radiophonique pensé en termes de « places »[213] (celle des locuteurs, celle des travailleurs du sons, et celle des auditeurs) et d'effets, que nous avons rassemblés sous la notion d'authenticité.

À la lumière de ces explications, nous souhaitons évaluer la prégnance de la voix dans la construction d'une « pragmatique radiophonique » auprès des travailleurs du son, qui l'emploient au quotidien (a) et des professionnels de la publicité et du marketing radio que nous avons interrogés (b). Nous introduisons cette séparation dans le traitement des entretiens entre les personnes travaillant le son directement et celles qui l'emploient de manière plus extérieure dans leurs métiers, car nous supposons que les rapports à la matière sonore sont distincts puisque mobilisés différemment dans l'exercice de leurs professions.

a) La voix, composante d'une pragmatique radiophonique chez les travailleurs du son.

Comme la finesse et la proximité avec les équipes antennes du travail réalisé par Hervé Glévarec laissait le présager, nous avons constaté au cours de nos entretiens avec les travailleurs du son que la voix est en effet au fondement de la pragmatique radiophonique. Elle transparaît dans les deux dimensions de cette notion : la vision du monde radiophonique, et sa mobilisation dans l'exercice des métiers.

Alban Rivoal, producteur à Fun Radio, considère la voix comme la « matière première » de son métier : « La voix en radio, c'est la matière première. On ne fait rien sans la voix, et tout le reste c'est ce qui va la mettre en valeur ». Ce matériau, les travailleurs du son le dissèquent et l'altèrent : « On découpe, on analyse, on va chercher énormément de choses... ». (…) Ça fait treize ans que je travaille en radio, et je pense que j'en aurai jamais fait le tour. (…) C'est ce qui nous attire là-dedans. On sculpte, on mixe, on fait ce qu'on veut avec le son »[214]. Nous pourrions lier cette perception de la voix comme un matériau aux notions de « dynéma » (développée par Pierre Schaeffer) et de celle

---

[212] ibid., p.15.

[213] Antoine, F., et. alii, *Analyser la Radio. Méthodes et Mises en Pratique*, éd. De Boeck, 2016, p.83.

[214] Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp. 154-158.

de « tailleur de sons », employée par Yann Paranthoën, réalisateur de documentaires pour Radio France aujourd'hui défunt, dans son autobiographie[215]. Pour exemple, Pierre Schaeffer emploie une métaphore chirurgicale pour décrire la transformation de la voix et du son en radio :

> « Sur la table d'opération, un son de cloche, de guitare, de piano [ou, une voix, une parole, un dire], gravé sur bande magnétique. D'un ciseau expert, quelqu'un vient de lui sectionner la tête ou les pattes. Le voici en trois morceaux. On les sépare par des "blancs". On recolle, on écoute. Voilà une dissection sonore que seule l'invention du magnétophone a rendue possible »[216].

En ce qu'elle est un matériau au référent sensible, la voix suppose tout d'abord une perception intuitive chez les personnes en maniant l'écriture comme Rafaële Dhuicque-Filloux, qui crée des spots publicitaires radio : « Mon métier il est très intuitif, très instinctif. Moi je le gère comme ça, avec mon ressenti, mes émotions, mes connaissances, ma petite culture, la perception du monde que j'ai autour de moi »[217]. La voix est par extension pensée et vécue comme le vecteur de nombreux effets sur les auditeurs — « Aujourd'hui ce que j'adore dans le fait de travailler en radio, c'est qu'on se demande sans cesse « qu'est ce que la voix provoque chez les gens ? » »[218] — , que nous avons identifiés précédemment et que les travailleurs du son viennent compléter et préciser. L'effet d'intimité est également interprété en termes de connivence entre l'animateur et l'auditeur par Julien Le Ny, producteur à Fun Radio, grâce à l'anonymat que permet ce média : « « Il y a un anonymat dans la radio qui pour moi doit se cultiver mais qui n'est plus d'actualité aujourd'hui »[219]. Ce dernier perçoit dans les effets de la voix de radio un mystère, qui vient compléter ce que nous avons qualifié d'effet d'imaginaire ou d'évocation que l'ajout d'image dénature :

> « Quand j'ai eu envie de faire de la radio au départ, il n'y avait pas les vidéos, les webcams... Ce n'était pas du tout mis en images. Moi je voulais savoir ce qu'il se passait pendant les disques. Je voulais savoir de quoi parlaient les animateurs, quelle était la vie d'une radio parce qu'il y avait un côté un peu mystérieux. Ce

---

215 Paranthoën, Y., *Propos d'un Tailleur de Sons*, éd. Phonurgia Nova, 2002.

216 Schaeffer, P., « Le laboureur et ses enfants », *op. cit.*, p. 225, §36, in Le Tinnier, F., « Dynamique de la voix enregistrée selon Pierre Schaeffer (1941-1963) », in *Entrelacs*, 2014. Disponible à l'adresse suivante : https://journals.openedition.org/entrelacs/621

217 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

218 Entretien avec Vladimir Bobrow, régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 5, pp. 161-163.

219 Entretien avec Julien Le Ny, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 2, pp.148-153.

> côté mystérieux a complètement disparu aujourd'hui, ce que personnellement je regrette »[220].

Évoquant la radio de son enfance, celle qui lui a donné sa vocation, le producteur mentionne l'objet qui en permettait l'écoute, le transistor. Contrairement aux enceintes connectées, qui répondent à un ordre de leur utilisateur, cet objet était l'occasion d'un voyage sur les ondes radiophoniques, naviguant de station en station sans toujours savoir ce que l'on entendait — une sérendipité qui a aujourd'hui elle aussi disparu lorsque la radio est écoutée en ligne :

> « Maintenant on sait tout le temps ce qu'on écoute, pourquoi on l'écoute, et on ne va plus découvrir la bande FM, sauf parfois quand on ne capte pas ses radios habituelles et qu'on va se balader sur des radios de province qu'on ne connaît pas et qu'on va découvrir par hasard. Mais ça c'est très rare maintenant »[221].

Si la voix en tant que matériau est transformée par son transport dans les ondes radiophoniques, l'application d'un format de station joue aussi un rôle dans son altération. Cette standardisation des modes d'expression est particulièrement audible pour les radios musicales, dont les *speaks* animateurs sont très courts et voués à lancer les prochains morceaux. Comment restituer le grain de voix, le laisser s'installer dans le temps long pour figurer les détails qui le composent ? Ce travail du détail est particulièrement présent dans le montage pour Hervé Glévarec, qui se réapproprie la notion d' « effet d'humanité » d'Albert Piette. Dans son *Ethnographie de l'Action*[222], l'anthropologue belge pense les détails de l'élocution comme périphériques à l'action principale, apportant un ancrage humain à l'énonciation d'un message, à la manière du second encodage théorisé par Iván Fónagy[223]. Pour Hervé Glévarec, les détails du grain de la voix sont au contraire maniés intentionnellement par les professionnels de la radio, afin d'ancrer les énonciations des locuteurs dans des corps :

> « A l'instar des salutations à l'antenne, qu'ils utilisent pour mettre en présence, les professionnels de radio renversent ici aussi l'ordre habituel : tandis que les détails produits dans une situation (dans un cadre de l'expérience) signifient l'humanité sur un « mode mineur de réalité », c'est grâce à la réintroduction de détails que les gens de radio, par un chemin inverse, produisent un effet d'humanité. De

---

220 ibid.

221 ibid.

222 Piette, A., *Ethnographie de l'Action. L'Observation des Détails*, éd. Métailié, 1996.

223 Fónagy, I., *La Vive Voix. Essais de Psycho-phonétique*, éd. Payot, 1983, p.12.

> contrainte interactionnelle, le détail devient ici ressource pour produire un effet d'humanité. *A contrario*, les professionnels de radio nous montrent que l'effet d'humanité est construit et que le détail est utilisé à propos, sinon stratégiquement. Le montage, comme travail *a posteriori*, en est la réalisation manifeste »[224].

Plus largement, Alban Rivoal ajoute à ce souci du détail locutoire en contexte radiophonique l'empreinte de la personnalité des animateurs dans les tonalités de leurs voix et les manières qu'ils ont d'aborder leurs *speaks*, donnant corps au propos. Le résultat : une sincérité ressentie par les auditeurs à l'écoute de l'antenne :

> « Le secret c'est de mettre de la vie. Les animateurs ont leurs petits secrets, mais leurs expressions sont les reflets de leurs vies (...). Je pense que le secret c'est de partager ce que tu ressens (…). Ça ne marche pas si tu n'es pas sincère avec les gens (…). Les *talks* de radio, si ce n'est pas la vie, ça n'intéresse personne. Du moins c'est mon point de vue »[225].

La problématique de la figuration d'une personnalité, d'une incarnation vivante d'un matériau vocal au référent sensible, est par conséquent au cœur d'une vision du monde articulée autour du son que les professionnels de la radio emploient dans la structuration de leur objet. Patrick Charaudeau parle ainsi du message radiophonique en tant que « produit fini », qui suppose à chaque étape de sa conception de garder une double dimension à l'esprit : les modalités de son « circuit de fabrication », et la mise en œuvre d'un « système de représentation de la pratique professionnelle »[226], que nous comprenons dans la notion de pragmatique radiophonique. Ces dimensions ne sont pas fixes. La pragmatique radiophonique, tant elle réunit d'acteurs et de compétences, induit régulièrement une « organisation par projets »[227], qui suppose d'en réactualiser les modalités à chaque nouveau départ. Les modalités de structuration de l'objet radiophonique sont perceptibles dans deux points évoqués par plusieurs des personnes que nous avons interrogées : l'écriture des messages radio et la sélection de comédiens pour les faire vivre à l'antenne.

Le travail de l'écriture radiophonique répond à une problématique propre, que nous avons qualifiée d'incarnation du matériau vocal, à l'œuvre dans plusieurs étapes de la conception du

---

224 Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001, p.156.

225 Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp. 154-158.

226 Charaudeau, P., « Introduction : Problèmes d'Analyse des Médias », in Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984, p.6.

227 Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001, p.12.

message. « (…) Le style du texte écrit destiné à être lu au micro est différent de celui du texte écrit ; il doit comprendre des phrases courtes, des mots concrets facilement compréhensibles. Ainsi du texte lu à la parole spontanée imprévue, il y a toute une gamme passant par le texte parlé, le débat préparé, la fausse spontanéité… »[228], qui permettent d'incarner le propos à l'oral. Vladimir Bobrow, régisseur-monteur à la station de radio RTL, explique son appréhension de l'écriture comme un maniement musical des signes linguistiques :

> « On parle à ce moment-là de fluidité de la voix, de ton, et de débit. On respecte ces trois qualités naturelles de la voix. Que ce soit pour nous, les régisseurs, ou pour les Speakerines qui vont lire nos textes — c'est hyper important de comprendre le débit d'une voix. Les respirations, les *hmmm*, les *euuuhh*... Ils permettent au moment du montage de garder l'état naturel de la voix (...). Notre enjeu c'est de conserver l'identité vocale et l'échange qui a lieu. Dans notre écriture, on utilise des onomatopées, on pense à la ponctuation selon qu'on écrit une phrase déclarative, affirmative, interrogative ou exclamative car les modes de lecture ne sont pas les mêmes »[229].

L'efficacité communicationnelle d'un message diffusé à l'antenne va jusqu'à faire l'objet de règles, instaurées dans le cas du groupe radio RTL par le Laboratoire Sonore, un service marketing visant à établir des standards d'efficacité. Le Laboratoire est à l'origine de règles que les personnes en charge de l'écriture s'efforcent de respecter. Pour exemple, un message comprenant plus de trois mots par seconde à l'oral sera *a priori* mal mémorisé par les publics. Cela ne peut cependant interdire d'en déroger, dans un souci d'originalité et de rupture que ne peut apporter que la création :

> « Le Laboratoire Sonore il fait un super travail en essayant de condenser tout ça dans quelque chose qui a été marketé et intellectualisé, mais par essence la création ne doit rien s'interdire. Moins il y a de barrières, mieux c'est. Après on ne peut évidemment pas faire n'importe quoi, ça reste un travail qui est maîtrisé parce qu'encore une fois, on ne fait pas de l'artistique, on fait de la pub donc on a quand même un cadre parce qu'il y a une obligation de résultat sur l'efficacité »[230].

L'incarnation du message ne s'arrête pas à l'écriture, qui doit entrer en cohérence avec la voix utilisée pour son enregistrement : « La base c'est d'enregistrer une voix. Une fois qu'on a une

---

228 Tudesq, A.-J., « Les conditions de production et d'écoute : leurs incidences sur le discours radiophonique », in Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984, p.15.

229 Entretien avec Vladimir Bobrow, régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 5, pp. 161-163.

230 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

voix et un texte on doit lui donner — comme une pièce de théâtre ou un film — une intention, une couleur et il doit y avoir un jeu »[231]. Le casting est un élément essentiel dans l'élaboration d'un message diffusé en radio, au même titre que l'écriture. Quelle sonorité vocale pour quel type de message ? Quel agencement verbal ? Le texte est la partition, et le comédien son interprète. Ils sont en cela indissociables, comme l'explique Rafaële Dhuicque-Filloux :

> « C'est le bon mot, au bon endroit, interprété par la bonne personne. On a la chance d'avoir une langue avec énormément de vocabulaire et de possibilités pour retranscrire par écrit, et ensuite à l'oral des émotions. Un petit mot, un petit adjectif, une petite respiration, un petit rire, un soupir... Une intention suffit à retranscrire une émotion, renforcée par la qualité d'interprétation du comédien (…). C'est joli parce qu'il y aura une harmonie entre la musique, les mots et les voix ; et ce sera efficace parce qu'on aura utilisé la bonne musique, les bons mots et la bonne voix »[232].

En outre, le travail de création des messages radio s'effectue dans le contexte de la station qu'il incarne. Une couleur d'antenne est vouée à être retranscrite dans chacune de ses composantes. Pour exemple, Vladimir Bobrow s'emploie à retranscrire la proximité que veut véhiculer RTL, comme une citation que nous avons utilisée plus tôt dans notre recherche l'illustre : « Aujourd'hui RTL table en termes de couleur d'antenne sur quelque chose de très proche des gens. La proximité est le maître mot : ça reste toujours la radio de la famille, transgénérationnelle »[233].

b) La voix, composante de la compréhension du travail radiophonique chez les professions publicitaires et marketing.

Nous substituons à la notion de « pragmatique radiophonique » celle de compréhension du travail en radio comme transformation de la voix pour les professions publicitaires et marketing, pour les raisons évoquées plus tôt. En effet, si ces personnes travaillent à et pour la radio, elles ne manient pas le matériau vocal au quotidien. Leur point de vue sur le contexte dans lequel leurs métiers s'insèrent provient par conséquent davantage d'une perception du média en tant qu'auditeurs avertis qu'en tant que créateurs du flux radiophonique.

---

231 Entretien avec Julien Le Ny, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 2, pp.148-153.

232 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

233 Entretien avec Vladimir Bobrow, régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 5, pp. 161-163.

Les effets perçus par les membres de ces professions correspondent à ceux que nous avons identifiés dans la première partie de notre recherche, avec quelques ajustements. Jérémy Valet, co-directeur de M6 Unlimited, voit en la radio un triptyque d'effets : l'évocation, que nous avons identifiée, ainsi que la proximité et la crédibilité, que nous interprétons davantage comme des résultantes de l'authenticité du média radiophonique, prolongée par sa fonction d'accompagnement structurant des habitudes d'écoute. Guillaume Duhamel concentre quant à lui son propos sur l'effet d'imaginaire de la radio, provenant du pouvoir du son, et argue qu'il a de puissant qu'il permet l'appropriation du matériau vocal par celui qui l'écoute :

> « Imagine un bord de mer, avec le bruit du port, des oiseaux, des vagues, éventuellement de gens qui marchent, du vent qui passe dans les feuilles d'arbre... Tu imagines une scène, un espace. Ça fait travailler l'imaginaire, et il n'y a rien de plus fabuleux pour l'esprit que d'imaginer les choses. Parce que tu vas te les approprier et créer les images en fonction de la perception que tu as du son »[234].

Le concepteur-rédacteur OPS fait le lien entre ce pouvoir évocateur et les émotions ressenties par l'auditeur — qui proviennent selon lui du référent sensible que traduisent ses sonorités :

> « Une voix va t'évoquer quelque chose, que tu vas imaginer. Si tu sens que la personne a une blessure… Par exemple, un poilu qui revient de la guerre avec le visage défiguré va parler d'une manière où tu vas imaginer ce visage, la souffrance qu'il doit vivre, et tout ça agit sur les émotions. Ça travaille sur ton imagination, et ça génère une émotion »[235].

Pierre Laurent, chef de produit marketing en charge du Laboratoire Sonore, ajoute deux dimensions au pouvoir du son. Premièrement, l'atmosphère qu'il instaure dans notre environnement — condition de possibilité de la fonction d'accompagnement de la radio et qualité qu'il partage avec le phénomène d'immersion virtuelle : « Le son c'est quelque chose qui nous accompagne au quotidien, et évoque grâce à cela plein de choses. Il est dans notre inconscient et rythme notre quotidien, entre dans notre intimité, voire dans nos émotions »[236]. Deuxièmement, son pouvoir d'ancrage dans la mémoire provenant de l'émotion qu'il procure, qui est cœur des logiques de répétions des messages en radio :

---

234 Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

235 ibid.

236 Entretien avec Pierre Laurent, Chef de produit chargé de la création et de l'analyse sonore au groupe radio RTL, mené par Clément Mellouet. Annexe 12, pp.187-191.

> « (...) le son de la cloche d'école. Il va réveiller des souvenirs en chacun de nous, parce qu'il a rythmé notre enfance dix ou quinze années durant. Il va peut-être te rappeler un camarade, ou ta salle de classe. Dans ce cas, le son agit un peu comme une madeleine de Proust. Le son peut réveiller tout un pan de notre mémoire qu'on aurait omis »[237].

L'ajout d'image au contenu audio — ou radio filmée — est vécu par l'ensemble de ces personnes comme une dénaturation du matériau vocal, dont les effets se déploient au sein d'un dispositif acousmatique. Guillaume Duhamel apporte cependant une nuance en considérant que si l'usage d'image peut s'avérer inapproprié, il peut constituer un complètement tout à fait enrichissant à l'expérience d'écoute :

> « Après l'image peut compléter l'expérience de manière intéressante. Par exemple quand Fun Radio fait la couverture de Tomorrowland [festival de musique électronique en Belgique], tu es content de pouvoir regarder les interviews des DJs en vidéo. Autre exemple : quand tu as des happenings dans une émission de libre antenne, c'est sympa de voir la vidéo. Ou encore quand tu as des artistes qui viennent faire une session acoustique dans le « Drive » sur RTL2 [émission musicale diffusée de 16h à 19h en semaine], il y a une partie interview que tu as envie de voir, leur session en live aussi »[238].

Comme l'explique Jérémy Valet, l'ajout d'image remet en cause le fonctionnement de la radio et son effet d'évocation, car elle dévoile tous les rouages de son illusion : « Aujourd'hui, ce que changent les réseaux sociaux dans une émission de radio, au-delà du fait que ça vient amplifier ce qui marchait déjà très bien, à savoir l'instantanéité et l'esprit communautaire — surtout l'esprit communautaire —, c'est dans l'exercice du métier : tu ne peux plus mentir »[239].

Nicolas Dot, chef de produit marketing en charge d'une veille sur les enceintes connectées, fait une interprétation originale de ce que nous avons qualifié d'authenticité du média radio. Pour lui, la radio demeure authentique car l'usage acousmatique de la voix est moins superficiel que la manipulation des images caractéristique de la virtualité numérique :

> « La possibilité d'uniquement conserver l'audio dans notre perception permet de s'attarder sur ce qui est dit, et d'avoir un jugement sur ce qui est dit et pas que ce

---

[237] ibid.

[238] Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

[239] Entretien avec Jérémy Valet, co-directeur de M6 Unlimited, mené par Clément Mellouet. Annexe 14, pp.194-204.

qui est vu, dans une ère du paraître, de la communication et des *fake news*. Dans ce contexte la radio a une pureté. On peut manier le discours pour manipuler mais c'est beaucoup moins facile à faire qu'avec des images »[240].

Le matériau vocal, déployé dans un dispositif acousmatique, est par conséquent un élément constitutif de la vision du monde que les professionnels de la publicité et du marketing radio emploient au quotidien. Le direct est également perçu comme nécessaire pour laisser la spontanéité de la voix s'exprimer, toujours en référence à l'authenticité du média : « Pour moi, l'essence même de la radio c'est le direct. Quelque soit le média — un Facebook Live, une émission TV, une émission de radio —, la sincérité, l'authenticité n'est possible qu'en live. L'imprévu, ça donne une dimension différente »[241], explique Guillaume Duhamel.

Les travailleurs du son et les personnes occupant des postes à la régie publicitaire du groupe radio RTL partagent une perception de la voix comme élément central des mondes radiophoniques qui encadrent leurs professions, et que les premiers mettent directement en œuvre au sein d'une pragmatique radiophonique. Le premier aspect de notre analyse de ces neuf entretiens semi-directifs est ainsi vérifié, laissant la place à notre prochaine préoccupation : les enceintes connectées, qui constituent le déplacement du matériau vocal dans une source artificielle, sont-elles perçues comme un indice de redéfinition plus large du média radio dans l'économie numérique ?

2. La perception des enceintes connectées comme un objet de redéfinition des métiers de radio.

Notre étude a entrepris d'analyser la radio sous le prisme des enceintes connectées, qui créent un champ concurrentiel élargi à une diversité d'entreprises médiatiques régies par une logique de média global, et déplacent le matériau vocal d'une source corporelle, physique, vers une Intelligence Artificielle dont les principes de conceptions visent à effacer la trace de cette désincarnation. Les enceintes connectées constituent ainsi un indice d'une progressive artificialisation des pratiques médiatiques, tant du point de vue de la création — qui obéit à des mécanismes de personnalisation algorithmiques détenus par des GAFAM — que des usages de

240 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

241 Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

consommation — dans un Internet dit symbiotique, créant un continuum entre la réalité matérielle et la virtualité. La radio se situe au carrefour de ces mouvements, car son travail historique du matériau vocal lui confère à la fois une légitimité face à ses concurrents et une osmose avec l'interaction et l'immersion par exemple à l'œuvre sur les réseaux sociaux.

Dans ce contexte, nous souhaitons interroger la perception des professionnels de radio quant au statut de cet objet dans la redéfinition de leurs métiers. Plusieurs interprétations peuvent être faites de la place de radio vis-à-vis de cette technologie. Charles-Emmanuel Bon, ancien directeur du développement du pôle radio RTL, faisait en cela une déclaration polysémique dans un article des *Échos* : « Les enceintes connectées pourraient devenir le relais de la radio dans le futur »[242]. Cette fonction de relais peut tout autant signifier que les enceintes connectées sont un support au travers duquel la radio peut s'exprimer et ainsi diversifier ses audiences, qu'indiquer une fonte de la radio dans un audio numérique que nous avons précédemment décrit. Le relais s'interpréterait ici comme une nouvelle étape, la nouvelle réalité de la radio, dans laquelle les stations historiques s'affaiblissent quand d'autres acteurs émergent, et supposant de prendre en compte l'artificialisation des pratiques médiatiques au sein de la virtualité numérique dans la création radio (une automatisation de certaines fonctions et la nécessité de repenser la monétisation des audiences en étant les points d'ancrage).

La deuxième étape de notre analyse de nos neuf entretiens semi-directifs répondra à la même division entre travailleurs du son (a) et professions publicitaires et marketing (b) que la précédente. En effet, le statut de la création radiophonique est directement questionné par des technologies telles que les enceintes connectées, quand la publicité et le marketing intègrent ces évolutions de manière plus extérieure, moins corrélée à la voix. Pour exemple, la veille sur les enceintes connectées dont est chargé Nicolas Dot indique le souhait de son service d'ouvrir de nouveaux canaux de monétisation, tout comme la création de M6 Unlimited, structure proposant des solutions *crossmedia* (qui mobilise plusieurs médias du groupe M6, alliant télévision, radio et digital) que Jérémy Valet co-dirige fait montre d'une place d'ores et déjà conséquente accordée à la remédiation numérique de la radio, combinant ses forces et ses faiblesses à d'autres leviers d'action. Au sein de nos groupes d'entretiens, notre analyse s'articulera en deux points successifs : la perception de l'objet que sont les enceintes connectées, puis sa mise en rapport avec l'évolution projetée des métiers des personnes interrogées.

---

[242] *Les Échos*, Alcaraz, M., « Les médias parient sur les enceintes connectées », 13/06/2018, n°22715.

a) Les enceintes connectées, indice d'une redéfinition des métiers des travailleurs du son.

Lorsque nous avons interrogé des travailleurs du son sur les enceintes connectées, notre propos s'ancrait dans un contexte qui influençait certainement leurs réponses. Questionnés tour à tour sur leurs parcours, leur perception du média et de ses effets, ses évolutions les plus prégnantes telles que l'ajout d'image avant d'aboutir aux enceintes connectées et à l'artificialisation des pratiques médiatiques, nos interlocuteurs développaient leurs discours dans une perspective d'évolution de la radio. La quête de leur avis sur les enceintes connectées provoquait dès lors une réponse en lien avec la radio et ses changements. Julien Le Ny a ainsi d'emblée lié les enceintes connectées aux modes d'écoute de la radio, confirmant sans le vouloir ce que nous avons qualifié de désincarnation du matériau vocal, en amenuisant les émotions :

> « Je trouve que c'est de l'évolution technologie, ce n'est pas de la radio. Les émotions qu'on peut avoir en radio, on n'est pas prêts à les faire passer là-dessus. Pour moi on ne l'utilise pas de la même façon. C'est un peu comme un assistant à qui on va demander telle ou telle chose, qui va répondre à des questions... Pour moi c'est deux mondes différents. L'assistant permet de pouvoir écouter la radio si tu le souhaites, mais ça s'arrête là »[243].

Le producteur interprète la notion de relais comme une diversification des modes d'écoute, et les contenus dédiés aux enceintes connectées comme des contenus d'appel vers l'expérience de l'écoute du flux radiophonique :

> « C'est un relais, oui ; il faut prendre ces outils en considération, oui, en passant aussi par Internet. C'est pour ça que le formatage de certaines rubriques à la radio ne doit pas dépasser deux minutes, ou une minute trente parce que c'est quelque chose qui peut se consommer sur les réseaux sociaux, qui peut se partager, etc. La radio maintenant elle se partage aussi sur Internet. Mais ces choses-là servent à donner envie de venir écouter la radio dans son entièreté, le média radio »[244].

Cette opportunité de diversification des modes d'écoute est vécue comme le possible déploiement vers de nouveaux types de contenus radiophoniques par Alban Rivoal, lui aussi producteur à Fun Radio : « C'est effectivement une opportunité. On serait bêtes de ne pas sauter dessus. Pour une fois

---

243 Entretien avec Julien Le Ny, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 2, pp.148-153.

244 ibid.

qu'on a une technologie, sorte de radio à laquelle on peut parler — elle te répond [rires] —, où il y a une interaction... Ce serait dommage de passer à côté »[245].

Si la perception des enceintes connectées comme un enjeu pour la radio est effective, elle est cependant relativisée par la plupart des personnes que nous avons interrogées. En interne, les enceintes connectées sont aujourd'hui uniquement une préoccupation pour le service innovation du groupe, comme l'explique Sacha Ephraim, régisseur-monteur à RTL : « Ce n'est effectivement pas quelque chose dont on entend beaucoup parler pour l'instant. Ce n'est pas quelque chose sur lequel on communique, à part au service innovation du groupe, qui cherche à développer des contenus sans que cela ne touche le reste des services pour l'instant »[246]. Rafaële Dhuicque-Filloux établit quant à elle un lien avec le retour du vinyle pour caractériser les enceintes connectées en tant qu'effet de mode :

> « Ce n'est pas parce qu'on multiplie les supports qu'un support va en remplacer un autre. Je vais te donner un exemple. Quand j'étais jeune on écoutait des 45 tours et des 33 tours, et on avait des platines avec un bras, un diamant puis on a tous fourgué nos vinyles parce que le CD est arrivé, puis après on a fourgué tous nos CD parce que tout est téléchargeable en numérique maintenant sur les plateformes. Et qu'est ce qu'il se passe là ? Et ben les gamins se ruent sur les ventes de vinyles, on réédite des vinyles. Johnny Hallyday est mort il y a trois mois, on a réédité toute sa discographie en vinyle. Donc il y a des effets de mode, mais Johnny Hallyday on l'écoute peu importe le support. C'est un éternel recommencement, c'est la mode qui revient »[247].

Comme nous l'avons observé plus tôt, le marché des enceintes connectées demeure embryonnaire sur le territoire français, contrairement aux États-Unis où la technologie connaît une forte croissance depuis son apparition en 2014. L'enjeu de redéfinition des métiers de radio que les enceintes connectées pourraient constituer est par conséquent dépendant de son degré d'appropriation par les publics d'un territoire, amenant nos interlocuteurs à en minorer l'importance.

L'arrivée d'une technologie présentant les atours d'un transistor, qu'elle augmente de la possibilité d'une interaction directe avec le contenu est en revanche perçue comme l'indice d'une évolution de la radio, et *a fortiori* des métiers du son. L'enjeu le plus prégnant est celui de la

---

245 Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp. 154-158.

246 Entretien avec Sacha Ephraim, Régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 6, pp. 164-167.

247 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

personnalisation du contenu à chaque utilisateur, que la majorité des personnes que nous avons interrogées considèrent crucial, comme l'explique Alban Rivoal :

> « Ça a commencé avec les podcasts, mais je pense qu'avec les enceintes connectées on va avoir un programme de plus en plus personnalisé. On va basculer — on est en train de le faire petit à petit — vers ça. On aura toujours des rendez-vous, notamment le matin, mais l'ensemble sera de plus en plus personnalisé. Je ne sais pas comment ça va se développer, mais j'en suis persuadé parce que les technologies telles que ces enceintes passent leur temps à essayer de personnaliser. Elles écoutent nos conversations, en fonction de nos conversations on reçoit des publicités… »[248].

Les enceintes connectées permettent en effet de cibler et de monétiser le contenu de manière efficace, comme nous l'avons évoqué précédemment. S'ajoute à cet engrangement des données vocales des utilisateurs la capacité de géolocalisation, également à l'œuvre dans des applications mobiles.

Conformément à ce que nous avons relevé lors de notre analyse de l'offre des stations de radio françaises sur Amazon Echo, les travailleurs du son ont souligné la nécessité de créer des contenus dédiés aux enceintes connectées. Pour Rafaële Dhuicque-Filloux, ces contenus doivent avoir de spécifique qu'ils mobilisent la personnalisation permise par les algorithmes à des fins d'interaction directe avec l'utilisateur :

> « Je trouve que pour nos métiers, c'est un virage qu'il faut prendre. Indéniablement. Parce qu'effectivement pour une fois c'est le son qui est mis en avant. Ça va nécessiter — j'espère — de créer des contenus dédiés et pas simplement être un relais de flux, juste un canal de diffusion (…). Là où c'est top c'est qu'on peut encore plus personnaliser les prises de parole, et pratiquement aller chercher chaque personne chez elle. Te parler à toi ce sera différent que de me parler à moi, à ta mère ou à nos enfants »[249].

Vladimir Bobrow souligne cependant les disparités concrètes entre le travail du flux radiophonique et la création de contenus délinéarisés, dédiés aux enceintes connectées, notamment en termes de temporalités : « Et n'oublions pas qu'on est un média. Il se passe mille trucs en dix minutes, tout

---

248 Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp. 154-158.

249 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

change tout le temps. Je pense plus globalement qu'on ne peut pas normer, informatiser cet aspect vocal-là »[250].

L'ultime aspect partagé par l'ensemble des travailleurs du son avec lesquels nous nous sommes entretenus est la place de l'humain — de leurs métiers — dans une économie numérique déployant une progressive artificialisation des pratiques médiatiques. Nos questions sur la présence d'une voix robotique, qui encadre ou remplace celle de l'homme dans des contenus audio, sont ainsi mises en rapport au sein d'une réflexion sur la relation entre l'homme et la machine. Cet aspect a été abordé de deux manières. Premièrement, le référent sensible de la voix est considéré comme le propre de l'homme, seul être à pouvoir lui donner corps et engager des audiences dans son discours. Rafaëlle Dhuicque-Filloux l'exprime en ces termes :

> « Est ce qu'on arrive à artificialiser l'odorat ? Pour moi la voix c'est pareil. Est ce qu'on arrive à robotiser le toucher ? Visuellement, maintenant on arrive à te tromper. Avec la réalité augmentée on arrive à te tromper. En revanche, le toucher on n'arrive pas à le tromper (…). La voix, l'ouïe, ça n'ira jamais — je pense — jusqu'à cet extrême parce que dans la voix, il y a dans le rythme, le grain de voix, la couleur, la tessiture, la vibration, quelque chose qu'on peut certainement programmer avec une machine mais j'ai peine à croire qu'on va parvenir à tout rassembler pour le restituer dans une machine. C'est ce qui nous est propre. On est les seuls à parler. Les mammifères communiquent par l'odorat, par certains cris et des sons, mais le langage tel que nous humains on a développé qui passe par la voix mais aussi par le langage corporel, je ne suis pas certaine que ça puisse disparaître »[251].

Deuxièmement, la place de la création humaine — dans notre cas des savoir-faire radiophoniques — est appréciée comme la valeur ajoutée et l'avantage concurrentiel de la radio au sein de l'économie numérique, dont nous avons observé le rôle stratégique dans le déploiement de ces entreprises médiatiques sur les enceintes connectées. La fonction de programmation identifiée par Hervé Glévarec serait par conséquent la plus grande force de la radio, dans un champ concurrentiel élargi qui tend à artificialiser les modes d'expression et d'accès aux contenus, comme l'explique Alban Rivoal :

> « [La radio] va évoluer, elle a déjà évolué. On ne peut pas rester dans son métier et regarder le train passer. Il faut être là au bon moment, prendre les bonnes

---

250 Entretien avec Vladimir Bobrow, régisseur-monteur à la station RTL, mené par Clément Mellouet. Annexe 5, pp. 161-163.

251 Entretien avec Rafaële Dhuicque-Filloux, fondatrice de la société de production Zéro Décibel, mené par Clément Mellouet. Annexe 8, pp.168-175.

décisions, parce qu'on tombe vite dans l'oubli. La règle c'est l'arène. La vie ou la mort. C'est très rapide (…). Pour retourner ce qu'on a dit sur les plateformes, peut-être que notre force dans l'avenir c'est qu'on n'est pas un algorithme »[252].

b) Les enceintes connectées, indice d'une redéfinition de la perception du média radio chez les professions publicitaires et marketing.

En concordance avec nos entretiens réalisés avec des travailleurs du son, les professionnels des pôles Opérations Publicitaires Spéciales (OPS) et Marketing ont été interrogés successivement sur leur perception des enceintes connectées en tant qu'objet, puis en tant qu'indice de redéfinition de la place du média radio au sein de l'économie numérique. Nous avons substitué ce second aspect à celui d'une redéfinition des métiers, car nous considérons que la perception du matériau vocal intègre la conception des pratiques de ces professions sans l'affecter directement.

Preuve de cette supposition, les enceintes connectées sont d'abord perçues par les membres de ces professions en tant qu'auditeurs, et non en tant que créateurs du flux radiophonique. Les enceintes connectées, capables d'enregistrer des conversations, constituent une intrusion des géants du numérique dans l'intimité de leurs utilisateurs qui suscite la méfiance de Guillaume Duhamel, concepteur-rédacteur OPS :

> « Sur l'enceinte elle-même, je trouve qu'elle contribue à la surveillance constante dont on fait l'objet dans nos sociétés. La vidéosurveillance, la reconnaissance faciale, ce genre de technologies permettent de créer des bases de données immenses. Avec ces objets, on construit de la *data*. L'enceinte connectée, ça fait que chez toi, tu as un objet qui écoute tout ce qui se passe, qui emmagasine des informations. Cet objet m'écoute, et je ne sais pas comment la *data* est traitée derrière. Il y a un danger. Je pense qu'on n'est pas assez informés sur ce risque »[253].

Nicolas Dot, en charge d'une veille sur les enceintes connectées, insère ce danger pour la vie privée dans les enjeux de monétisation algorithmiques que nous avons préalablement esquissés :

---

252 Entretien avec Alban Rivoal, Chargé de Production à Fun Radio, mené par Clément Mellouet. Annexe 3, pp. 154-158.

253 Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

> « Je pense que les enceintes connectées font partie d'un écosystème de développement des GAFA qui crée une sorte de paranoïa ou d'inquiétude pour la vie privée qui peut être compréhensible. Les enceintes connectées, d'un point de vue commercial et publicitaire, permettent de récupérer un maximum de données, surtout vocales afin de pouvoir pousser des promotions et, au lieu de faire un achat par le *search*, faire un achat par la voix. Cela permettrait de faciliter encore l'acte d'achat en le raccourcissant et en le ciblant davantage, en fonction des habitudes d'écoute et des préférences d'utilisation de chacun »[254].

Résultante de leur fonctionnement algorithmique, les enceintes connectées représentent pour le chef de produit marketing un enjeu à la fois éditorial et publicitaire pour la radio. La fonction de relais évoquée par Charles-Emmanuel Bon, ancien directeur du développement du groupe RTL, est ici interprétée en un cercle vertueux, où le ciblage des contenus radiophoniques permet une monétisation centrée sur des publics au fort potentiel commercial :

> « Le relais, je le comprends en deux axes : l'axe publicitaire et l'axe éditorial. Je commence par l'axe éditorial. Les enceintes connectées relais de la radio, c'est d'abord comme un nouveau support d'écoute, un objet qui est installé dans ton foyer comme un transistor mais avec des fonctionnalités plus importantes. Elles permettent donc de relayer l'information ou le divertissement, et après servent de relais pour l'approfondissement du sujet, pour se renseigner. Le second relais est pour moi d'ordre publicitaire, hautement stratégique car c'est ce grâce à quoi la radio est viable et peut proposer une diversité et une richesse éditoriale importante. La publicité va avec les enceintes connectées être amplifiée et va pouvoir générer un acte d'achat qui va permettre à terme d'avoir des relais publicitaires plus conséquents, installant un cercle vertueux avec des contenus éditoriaux encore plus développés »[255].

L'artificialisation des pratiques médiatiques dont les enceintes connectées constituent un indice est un autre point récurrent des entretiens que nous avons réalisés. Dans un parallèle entre la Réalité Virtuelle et les enceintes connectées, Guillaume Duhamel étaye un exemple qui résonne avec le principe de la *uncanny valley* (« la vallée dérangeante ») que nous avons précédemment développé. Selon lui, la projection hors du monde virtuel nous fait éprouver son caractère fétiche, faussé — un choc émotionnel aux attraits monstrueux, signe d'une artificialité rejetée par l'être humain. Le concepteur-rédacteur croit cependant en la force de l'habitude, qui peu à peu — dans une situation simultanée de progression de la technologie —, permettra son adoption par les publics,

---

254 Entretien avec Nicolas Dot, chef de produit marketing au groupe radio RTL, mené par Clément Mellouet. Annexe 10, pp.178-184.

255 ibid.

et contredisant le principe du *uncanny wall* (« le mur dérangeant ») avancé par certains chercheurs comme un frein inéluctable à l'entrée de la technologie dans nos existences :

> « Ça me fait penser à un autre exemple : la réalité virtuelle. J'ai récemment testé ce produit, auquel on ne répond pas tous de la même manière selon notre cerveau et notre personnalité. Quand j'ai fait le test, on mettait le casque, on était à New York et on montait dans un ascenseur. Dans cette scène que l'on te fait vivre, tu peux tourner la tête, comme dans un ascenseur panoramique, vitré. Quand tu arrives tout en haut, il y a une poutre et le vide. Tu as vraiment l'impression d'y être. On te demande de monter sur la poutre. Toutes les personnes qui ont fait le test tremblaient à ce moment-là, alors qu'on avait juste un casque et qu'on marchait sur le sol. Tu peux aller dans l'immersion encore plus loin en mettant un tasseau de bois au sol pour récréer la poutrelle d'ailleurs. L'impression de vrai est en tout cas très forte. Une animatrice nous accompagne. Au moment de monter sur la poutre, elle te prend la main. Ça te rassure un petit peu, même si on perd soudain en cohérence avec ce qu'on visualise dans le casque. A la fin, pour aller au bout de l'expérience, tu dois sauter. Beaucoup ont juste fait un pas en avant. Moi j'étais tellement dedans que j'ai vraiment sauté. Ça a déconnecté le casque et je me suis vautré par terre. A ce moment-là, j'ai eu ce que tu as décrit. Le choc émotionnel de me rendre compte que tout ça est faux. Imaginons que tu parles avec quelqu'un dans ta cuisine, de manière très fluide, et qu'en fait il s'agisse d'une enceinte connectée. Ça te fait forcément quelque chose d'assez bizarre. Peut-être que les premières fois avoir un échange avec une machine va te choquer, mais qu'après tu vas t'habituer. L'être humain s'habitue. A un moment donné, on va s'habituer à interagir avec des outils, parce qu'en plus c'est pratique, ça simplifie nos tâches. De la même manière que les portables sont entrés dans nos vies. Les expériences changent avec la technologie. Dans chaque cas, il y a du négatif comme il y a du positif »[256].

L'exemple est cependant contrebalancé par son auteur, qui nuance cette progressive imprégnation de la technologie dans la vie humaine par la place de l'humain dans la création médiatique. Nous avons en effet observé que la voix, en tant qu'incarnation des mouvements d'un corps, est altérée par le déplacement de sa source dans une Intelligence Artificielle. Outre l'importance de la fonction de programmation de la radio relevée par Alban Rivoal en contexte algorithmique, le grain de la voix est ce qui donne un ancrage, une matérialité et une complexité au propos :

> « Le point fort de la radio sur les enceintes connectées comme on les connaît actuellement, ce sont d'abord ses voix. La voix artificielle est de plus en plus réussie, mais ce qui compte c'est l'humain derrière la voix. En plus de ça, la radio accompagne, comme cette enceinte, dans le quotidien. Et la radio sait travailler

[256] Entretien avec Guillaume Duhamel, Concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL, mené par Clément Mellouet. Annexe 16, pp.207-213.

> l'audio. Elle est d'emblée plus fluide et naturelle, et contrairement à l'Intelligence Artificielle peut construire des contenus complexes. L'enjeu sera aussi de personnaliser les contenus de la radio grâce à ce type de technologies. La radio, elle est composée de gens dont c'est le métier — des régisseurs, des monteurs, des réalisateurs, des animateurs qui ont une signature… »[257].

La perception des enjeux auxquels fait face le média radio au sein de l'économie numérique de la voix est par conséquent aiguë. Elle se conjugue à des considérations sur la place de la radio, en termes de stratégies de positionnement, sur les enceintes connectées. Jérémy Valet, co-directeur de M6 Unlimited, ajoute à une problématique de juste équilibre entre les offres linéaires et délinéarisées pour la cohérence d'une image de marque — que nous avons développée dans la seconde partie de notre recherche — la description d'un champ concurrentiel où le privilège de la radio sur les supports audio lui est dénié par ses concurrents. Ce phénomène est accentué par la défaillance de notre média quant à son appropriation des technologies, qui contraint la création de contenus dédiés aux enceintes connectées :

> « *[Pensez-vous que la radio, média de la voix et du son, ait sur ce support une légitimité face à ses concurrents, qui n'ont pas les mêmes capacités en termes de production sonore ?]*. C'est la conviction que l'on a. Maintenant tu vas avoir d'autres acteurs du marché qui te diront "alors nous on n'est pas experts de l'audio mais on est experts du digital et des nouveaux *devices*, et on sait développer des technologies qui nous permettent de faire des choses que les gens du média historique de la radio ne savent pas faire". De toute façon c'est toujours pareil : quand il y a un nouvel item qui est lancé, il y a plusieurs personnes qui revendiquent être les mieux placées pour répondre à la problématique, et chacun a ses avantages. Evidemment les avantages de la radio c'est qu'on maîtrise le média qui passe uniquement par le canal audio. Maintenant ce qu'on maîtrise certainement moins que d'autres acteurs, c'est toute la flexibilité ou en tout cas toute la modernité dans l'offre techno, tout la dimension techno qu'on pourrait apporter à notre offre »[258].

Cette défaillance technologique, la radio la partage avec nombre d'autres médias historiques. L'aide fournie par Binge Audio, studio français de podcasts natifs, au *Parisien* pour créer des contenus sur Amazon Echo en est un exemple. Les stations de radio composent ou s'insèrent dans de grands groupes, comme la fusion entre les radios RTL, RTL2 et Fun Radio avec M6 en atteste. Lorsque interrogés sur leur risque de sclérose face à des acteurs plus petits mais empreints de flexibilité et de dynamisme, nos interlocuteurs voient plutôt en le poids des groupes de

---

257 ibid.

258 Entretien avec Jérémy Valet, co-directeur de M6 Unlimited, mené par Clément Mellouet. Annexe 14, pp.194-204.

radio un avantage concurrentiel. Pierre Laurent, chef de produit marketing en charge du Laboratoire Sonore, rappelle ainsi l'ancrage de la radio dans le quotidien des Français :

> « En France, quarante million de personnes écoutent la radio chaque jour. Ça donne une assise tellement puissante qu'il faudra du temps pour détrôner ça. Les assistants vocaux aujourd'hui, c'est très émergent. On en parle parce que ça fait le buzz. Il y a Apple, Amazon, Google... On sait que ce sont des entreprises qui sont très puissantes donc on se doute que ça va marcher. Mais mettre les chiffres d'audiences de la radio en rapport à cela permet de se rappeler que ce média reste ancré dans le quotidien des Français »[259].

Si la radio possède une avance dans le développement de contenus audio, la remédiation de son matériau sur le numérique suppose une conduite du changement afin de mobiliser l'ensemble des savoir-faire des métiers du son dans la construction d'une offre dédiée aux enceintes connectées et aux plateformes d'offres délinéarisées. Jérémy Valet applique cette considération à son groupe :

> « Peut-être qu'il fallait qu'on attende d'être dans le groupe M6. On a les compétences audio, M6 a les compétences liées à la techno et à l'innovation, et peut-être qu'en mixant les deux expertises on sera les mieux placés pour réussir à mettre en place un mouvement entre la production pour la radio et la production de contenus RTL, RTL2 et Fun Radio pour les enceintes connectées »[260].

Pierre Laurent résume cette tension entre le poids historique des stations de radio et le dynamisme que semble requérir le numérique avec les notions de tradition et d'innovation, qui supposent de mêler la stratégie de média aux ressources humaines, dimension que nous ne pouvons traiter en profondeur dans notre analyse, laissant la porte ouverte à d'autres travaux dans ce domaine :

> « La technologie fait qu'on peut accéder à tout instantanément, et la radio doit être dans cet état d'esprit. Elle le fait avec les podcasts et les *replay*, qui prennent une ampleur grandissante. Même si la radio doit conserver son aspect traditionnel, où on sait qu'à 8h il y a le journal. La majorité de la population consomme aujourd'hui la radio de manière linéaire, comme l'ont fait nos parents et nos grands-parents »[261].

---

259 Entretien avec Pierre Laurent, Chef de produit chargé de la création et de l'analyse sonore au groupe radio RTL, mené par Clément Mellouet. Annexe 12, pp.187-191.

260 Entretien avec Jérémy Valet, co-directeur de M6 Unlimited, mené par Clément Mellouet. Annexe 14, pp.194-204.

261 Entretien avec Pierre Laurent, Chef de produit chargé de la création et de l'analyse sonore au groupe radio RTL, mené par Clément Mellouet. Annexe 12, pp.187-191.

Notre étude de neuf entretiens semi-directifs menés avec des travailleurs du son et des personnes occupant des professions publicitaires et marketing au groupe radio RTL nous a permis d'évaluer le statut des enceintes connectées en tant qu'indice d'une redéfinition de leurs métiers, articulé en deux points.

La prégnance de la voix comme élément de définition d'une « pragmatique radiophonique » est perceptible chez l'ensemble des personnes que nous avons interrogées, qui voient en les propriétés du matériau vocal un fondement des effets du média radio, et le mettent en œuvre dans leurs professions. Conformément à la théorisation de Hervé Glévarec, la voix est une forme que les travailleurs du son mettent en forme — elle est la condition et l'objet de leurs métiers. Les professionnels de la publicité et du marketing attribuent des effets au média radio qu'ils vont quant à eux mobiliser dans la conception de dispositifs d'Opérations Spéciales ou d'argumentaires pour des annonceurs, l'incidence indirecte de la voix dans leurs professions supposant de remplacer la notion de pragmatique par celle de compréhension de cette dimension dans leur travail.

La perception de enceintes connectées comme un objet de redéfinition des métiers de radio est en revanche plus nuancée. Cet enjeu ne semble pas avoir intégré l'ensemble des antennes du groupe radio RTL, se cantonnant à la cellule innovation. L'apparition de cette technologie, portée par des GAFAM, est considérée comme une traduction — que nos questions ont convertie en la perception d'un aboutissement, car la radio partage le matériau vocal avec ces enceintes — de logiques de personnalisation des contenus dont la radio doit s'emparer. L'état embryonnaire du marché des enceintes connectées — dont nous ne pouvons présager l'adoption par les publics en France — fait de notre objet un indice d'une redéfinition des métiers de la création médiatique, caractérisée par son artificialisation, sans qu'il ne soit corrélé à l'exercice des métiers de radio. Les personnes que nous avons interrogées ont toutes adopté un point de vue extérieur en abordant cet enjeu, focalisant la conversation sur les évolutions de la relation homme-machine davantage que sur leurs professions. Le travail du son, s'il intègre des évolutions technologiques, reste en effet inchangé par ces technologies. C'est peut-être davantage la manière d'aborder la diffusion des messages, qui devient plurielle, qui progressera dans le maniement du matériau vocal au quotidien.

En traitant l'appropriation de la voix par une multiplicité d'acteurs dans l'économie numérique, qui dépasse le marché des enceintes connectées, avant de la mettre en relation avec la perception que les personnes travaillant en radio ont de leurs métiers, nous avons entrepris d'évaluer la prégnance d'un enjeu encore en construction. Des médias divers organisent leurs stratégies vocales, à l'état de conception, et les professionnels de la radio ne vivent les évolutions

apportées par les enceintes connectées que comme peu corrélées aux pratiques de leurs métiers. Les enceintes connectées sont perçues comme un enjeu de diversification et de personnalisation des prises de parole grâce à l'interaction qu'elles proposent, sans pour autant engager une réflexion sur une redéfinition plus large des métiers de radio sur le numérique. Nous avons pourtant observé que les enceintes connectées mettent à jour une perte de spécificité du média radio, dont les fonctions et le matériau vocal sont accaparés par des acteurs qui lui sont extérieurs et dont elle subit la concurrence (dans le cas des autres entreprises médiatiques) ou la domination (notamment des concepteurs des technologies, à l'origine des enceintes connectées et des plateformes numériques). Dans ce contexte, la radio bénéficie d'un avantage concurrentiel provenant de son identité vocale et sonore, mais — surtout en France — semble peu prompte à investir pleinement l'aboutissement de l'intégration de l'audio dans l'environnement de la virtualité numérique que constituent les enceintes connectées. Si nous pensons que la réflexion des stations de radio sur les enceintes connectées doit être d'abord portée par leurs identités vocales et sonores, nous avons mis à jour la nécessité pour une entreprise médiatique de faire résonner son support avec une image de marque. C'est au sein de cette tension entre la signification du savoir-faire radiophonique et le déploiement d'une identité commerciale que les stations de radio doivent axer leurs stratégies de diversification sur les enceintes connectées. Nous allons tenter de décliner ce constat en une série de recommandations relatives au positionnement des stations de radio sur les enceintes connectées. À plusieurs reprises seront synthétisés des points évoqués au fil de notre réflexion, ici articulés dans une stricte perspective de stratégie de positionnement.

C. RADIO DÉCADENTE OU RADIO RENOUVELÉE ? PERSPECTIVES DE POSITIONNEMENT DES STATIONS HISTORIQUES SUR LES ENCEINTES CONNECTÉES.

1. Les stations de radio au sein d'un champ concurrentiel élargi : une tension entre l'identité vocale et l'image de marque.

Notre analyse du statut de la radio sur les enceintes connectées nous a permis d'établir l'existence d'un champ de création médiatique qui, peu investi par les stations, comporte autant de menaces que d'opportunités pour ce média. Tout comme notre recherche est marquée par sa démarche exploratoire, contrainte par une économie en cours de construction, le temps dont dispose la radio pour créer un réflexe vers son média sur ce nouveau support est limité. Le marché qu'ouvrent les enceintes connectées constitue une opportunité pour un vaste nombre de marques régies par la notion de média global de construire leurs voix propres, et de s'imposer. Au sein de cet environnement, les stations de radio disposent d'une avance sur laquelle elles peuvent miser : l'attachement de leur support d'origine, constitutif de leurs identités, au matériau vocal et au maniement de son authenticité. À cet avantage s'adjoint la nécessité pour les stations de radio de se remédier, dans une économie numérique qui prend le pas sur les revenus publicitaires générés par la vente d'espaces sur leurs flux linéaires. Un récent rapport publié en 2018 par le ministère de la culture français à propos des médias et de la publicité note un transfert de valeur significatif des revenus publicitaires des médias « historiques » (la télévision, la presse et la radio) vers Internet[262]. Menée en collaboration avec le Conseil Supérieur de l'Audiovisuel (CSA) et le cabinet de conseil Bearing Point, l'étude remarque qu'alors que les médias historiques avaient un quasi monopole des recettes publicitaires jusqu'en 2002, leur part de marché a décru à 65% en 2017, en voyant 35% accaparés par Internet[263]. Mise en rapport avec les autres supports, la radio apparaît comme le parent pauvre des dépenses en communication des annonceurs, qui s'élevaient à 800 millions d'euros en 2017 sur un total de 10,7 milliards d'euros investis en média[264]. Ce plus faible investissement des annonceurs, s'il est structurel et non récent, accentue selon nous l'intérêt pour les stations de se développer sur des supports d'écoute numérique en forte affinité avec leurs

---

[262] Ministère de la Culture, « Médias et Publicité en Ligne », 25/07/2018. Disponible à l'adresse suivante : http://www.culture.gouv.fr/Thematiques/Audiovisuel/Rapports-etudes/Etude-sur-les-Medias-et-la-publicite-en-ligne-transfert-de-valeur-et-nouvelles-pratiques2

[263] ibid., p.14.

[264] ibid., p.9.

identités et leurs modes de consommation, dont les enceintes connectées constituent un aboutissement. À enjeu s'ajoute celui concomitant d'un vieillissement des structures d'audiences des stations sur leurs flux linéaires, que nous avons évoqué plus tôt, faisant des enceintes connectées une opportunité d'autant plus vive de renouvellement.

Les stations de radio peuvent ainsi créer un réflexe chez les utilisateurs d'enceintes connectées de deux manières : en déployant un discours d'escorte de leurs contenus mettant en avant leur légitimité sur ce support, et en créant des contenus mobilisant leurs savoir-faire tout en en incorporant de nouveaux, leur permettant de mobiliser pleinement l'interaction directe apportée par cette technologie. Le travail d'un discours d'escorte, s'il mobilise une légitimité dans le travail de la voix, suppose une réflexion propre à chaque marque. La légitimé peut être exprimée de manière historique, qualitative ou en termes d'audiences. Pour exemple, RTL, qui forte de ses audiences linéaires se présente comme la « première radio de France », assoit sur l'ensemble de ses supports d'expression une légitimité pourtant peu prégnante sur le numérique, où seules « Les Grosses Têtes » et la pastille humoristique de Laurent Gerra sont beaucoup écoutées. Le discours d'escorte, pour être pleinement efficient, ne peut dès lors provenir que de la qualité des contenus qu'il incarne et représente. Cette qualité, si elle peut s'évaluer avec une diversité de critères parfois peu objectifs, doit s'adjoindre d'une cohérence avec les modes de consommation permis par les enceintes connectées. Si nous avons mis à jour que la voix partage avec la virtualité numérique des propriétés d'immersion et d'interaction, la première semble déjà acquise pour la radio quand la seconde nécessite un important travail.

Un investissement tout à fait efficient des enceintes connectées par les stations de radio historiques suivrait un parcours logique allant de la stratégie de contenus à leur encadrement marketing. La création de contenus dédiés aux enceintes connectées — ou du moins réadaptés à leur interaction immédiate avec l'utilisateur — suppose à la fois d'ajouter des compétences aux professionnels de radio et d'engager une réflexion sur la cohérence des offres linéaires et délinéarisées. Comme nous l'avons observé précédemment, l'audience de la radio demeure attachée à son flux linéaire, dont l'authenticité en situation d'accompagnement est accrue par un dispositif acousmatique — l'écoute numérique représentait 12,2% de son audience sur la vague janvier-mars 2018 en France[265]. L'offre audio proposée par les stations de radio sur le numérique se compose aujourd'hui majoritairement de contenus d'appel. Pour exemple, le flash de franceinfo sur Amazon

[265] *Offre Média*, « Les supports multimédias atteignent 12,2% du volume total d'écoute de la radio selon Médiamétrie », 18/05/2018. Disponible à l'adresse suivante : https://www.offremedia.com/les-supports-multimedias-atteignent-122-du-volume-total-decoute-de-la-radio-selon-mediametrie

Echo, réactualisé quatre fois par heure, est un chemin vers l'expérience de l'écoute en direct, au cœur de l'information ; les pastilles de radio filmée diffusées sur les réseaux sociaux sont des fragments appelant à l'écoute de la radio, où l'expérience est instantanée et authentique. Cette stratégie est tout à fait pertinente, puisque l'audience linéaire est la force principale de la radio. Un manque de cohérence entre les offres linéaires et délinéarisées fait en outre courir le risque à la station qui les incarne de perdre en unité aux oreilles de ses publics. Un investissement dans les modes d'écoute digitaux ne peut nier un existant radiophonique, source de ses revenus et de sa puissance. À la lumière de ces considérations, le caractère embryonnaire du marché des enceintes connectées constitue un frein au développement des stations de radio sur ce support, qui en outre manque d'une mesure d'audiences stable. Néanmoins, dans un contexte où ses revenus publicitaires s'y amenuisent, la conquête des enceintes connectées constitue pour la radio un levier attaché à la voix, même s'il en diffère sur plusieurs points, auxquels les savoir-faire radiophoniques doivent s'adapter. Nous allons dans notre prochaine partie traiter de cette intégration de nouvelles pratiques radiophoniques sous deux aspects saillants, laissant de côté une nécessaire conduite du changement davantage attachée aux ressources humaines que nous n'avons pas l'entière capacité de traiter ici : la conception de parcours d'utilisation adaptés à l'interaction permise par les enceintes connectées, et l'ajout de compétences dans le maniement des Intelligences Artificielles qui encadrent et amènent au contenu radiophonique.

2. L'artificialisation des pratiques médiatiques : incorporer un savoir-faire technologique dans la création radiophonique.

L'intégration de savoir-faire technologiques dans le processus de création radiophonique intervient à plusieurs étapes, que nous résumons en un double travail : un maniement de l'environnement algorithmique, et la construction de modes de relation ente l'utilisateur et le contenu propres aux enceintes connectées.

Dans un article consacré aux médias d'information sur les technologies vocales[266] (englobant ainsi tous les supports où sont présents des assistants vocaux, allant de l'enceinte au mobile) Trushar Barot, concepteur d'applications pour les services d'information de la British Broadcasting Corporation (BBC), propose une analyse compréhensive des enjeux auxquels font

[266] *Nieman Lab*, Barot, T., « The future of news is humans talking to machines », 18/09/2018. Disponible à l'adresse suivante : http://www.niemanlab.org/2017/09/the-future-of-news-is-humans-talking-to-machines/

face les acteurs présents sur ce support. Nous concentrons notre intérêt sur l'un des points abordés dans l'article : la distribution des contenus aux utilisateurs. Insérée dans une plateforme qui, si elle peut être en partie parcourue à l'aide d'un mobile, privilégie l'ouïe et la voix, comment l'application d'une entreprise médiatique peut-elle atteindre un utilisateur auquel l'assistant vocal ne propose qu'un nombre limité de choix ? En effet, contrairement à la liste de résultats que présente un moteur de recherche internet classique, Alexa, Google Assistant ou SIRI ne formulent qu'une réponse. Cette problématique gagne en complexité dans un environnement caractérisé par la concurrence élargie et la pluralité de manières possibles de formuler une requête que nous avons précédemment étudiées. Un premier élément de réponse consiste à nouer des partenariats avec ces plateformes, même si cette solution ne peut être systématique pour des raisons évidentes de rentabilité et de concurrence entre les marques médias. La courte fenêtre temporelle aujourd'hui encore ouverte aux entités qui souhaitent se positionner parmi les acteurs phares de ce marché est en outre contrainte par un phénomène de « loyauté accidentelle »[267], où l'utilisateur reviendra mécaniquement vers les marques qui ont su le conquérir en premier, leur décision étant portée par l'habitude et les suggestions des algorithmes. Afin d'affiner les recommandations des assistants vocaux, et d'affirmer la place d'un média au sein de leurs suggestions, Trushar Barot invoque le déploiement d'un système de classification des contenus qui outrepasserait une dimension thématique. Les traditionnels mots-clés apposés au bas d'un article sur le web se verraient enrichis par des *tags* comprenant également les situations de consommation auxquelles un contenu se prête le mieux : « Ces interfaces auront la capacité de détecter la situation et l'état d'esprit d'une personne à un moment donné, leur permettant de déterminer la manière dont elles interagissent avec la personne à ce moment précis »[268] *[« These devices will have the ability to detect a person's situation and well as their state of mind at a particular time, enabling them to determine how they interact with the person at that moment »]*. Si l'article fait d'abord référence aux différentes situations d'écoute d'une personne embarquant un assistant vocal dans un mobile, alors qu'une enceinte connectée reste dans un foyer, nous pouvons arguer que parmi les paramètres à prendre en compte, l'heure à laquelle est demandé le contenu représente une première information de taille. En effet, un contenu d'information demandé le matin ne pourra avoir la même durée ni la même structure narrative qu'un contenu de soirée, plus long et profond. L'avantage de la radio est ici immense, puisque les

---

[267] *Quartz*, Gerhart, A., « Alexa makes decision-making easier than ever—by making your choices for you », 10/07/2018. Disponible à l'adresse suivante : https://qz.com/1323884/alexa-makes-decision-making-easier-than-ever-by-making-your-choices-for-you/

[268] Traduit de l'Anglais, *Nieman Lab*, Barot, T., « The future of news is humans talking to machines », 18/09/2018. Disponible à l'adresse suivante : http://www.niemanlab.org/2017/09/the-future-of-news-is-humans-talking-to-machines/

stations font d'ores et déjà montre d'une capacité à proposer des programmes adaptés à leurs situations d'écoute, qui sont présupposées à partir des horaires de diffusion sur le flux en direct. Ainsi intervient le second travail que nous avons évoqué : la création de contenus dédiés à l'interaction directe avec l'utilisateur, qui diverge de l'écoute partiellement passive (car dotée d'une dimension sociale à l'antenne et en ligne) de la radio.

Dans une *masterclass* menée en juin 2018 par la cellule en charge de l'innovation du groupe Radio France[269], les enjeux propres à la création de contenus dédiés aux enceintes connectées étaient exposés de manière méthodique. Nous focalisons notre réflexion sur une partie des points évoqués lors de cette présentation, car nous avons évoqué les autres dans le précédent aspect de notre analyse. Sont mis en lumière les nécessaires travaux sur l'arborescence de l'*app* ou de la *skill* (nous l'avons vu, les terminologies diffèrent selon le développeur), et de l'identité de la marque qui s'exprime sur l'enceinte connectée.

Le premier peut être qualifié d'*user experience design*, relatif à la conception des expériences des utilisateurs. Contrairement à une application mobile, qui dispose d'un appui visuel permettant des aller-retours et une importante exploration de l'offre d'un média, le nombre de choix doit être réduit et la clarté de son utilisation amplifiée. Ce travail est encore à l'état d'expérimentation, expliquant sans doute que notre tableau comparatif des offres des stations de radio sur Amazon Echo en France[270] mette en lumière l'existence d'une pluralité de *skills* pour un seul et même média (notamment franceinfo, dotée de cinq *skills*, ou encore France Inter, qui en compte quatre). Rassembler l'ensemble des programmes proposés par la station sur une application aurait probablement constitué un espace trop laborieux pour être exploré, car supposant de l'utilisateur qu'il retienne une longue liste de propositions avant de faire son choix. Le nombre d'interactions doit par conséquent être réduit et simplifié dans l'accès au contenu au sein de l'application. À l'inverse, il doit être déployé dans le contenu dédié à l'enceinte connectée, proposant une expérience supplémentaire à celle de l'écoute d'un flux linéaire ou délinéarisé. Outre le flux en direct et des podcasts, il semble pertinent pour une station de proposer des formats de questions-réponses ou des narrations sur lesquelles l'auditeur peut avoir une influence directe dans leurs dénouements. Cela suppose d'allouer des ressources à un marché encore restreint, et par conséquent de prendre un risque plus ou moins contrôlé selon la taille de la structure et les compétences technologiques du groupe portant une station qui s'investit dans ce marché.

---

269 Radio France, Masterclass, « Smartspeakers / Assistants Vocaux : Comment se Faire Entendre ? », 11/06/2018. Disponible à l'adresse suivante : http://hyperradio.radiofrance.fr/innovation/inscrivez-a-masterclass-radio-france-station-f-8-juin/

270 Annexe 18, pp.220-222.

Le second point présentant un intérêt à notre formulation de recommandations révèle à nouveau un avantage de la radio sur ses concurrents : la nécessité de signifier l'ancrage du contenu dans l'offre d'une marque à l'identité singulière, avec la voix et le son pour seuls appuis. Les techniques d'habillage sonore et de rappel oral du nom de la station sont ici cruciales, mais déjà maîtrisées par les stations. L'ajustement effectué par le groupe Radio France pour le flash de franceinfo sur Amazon Echo est ainsi minime : un enregistrement d'une personne annonçant l'heure et la station (pour exemple, « il est 8h20, vous écoutez franceinfo ») a été effectué puis inséré en début de chaque rappel des titres, rattachant l'expression du journaliste à un flux en direct, crucial pour une station d'information en continu.

Les recommandations que nous avons formulées quant au positionnement des stations sur les enceintes connectées font état d'une redéfinition du média radio au sein de l'économie de la voix numérique. Concurrencée par des entreprises de presse, de télévision, ou issues d'Internet, et encadrée par des systèmes algorithmiques animés par des Intelligences Artificielles, la radio a selon nous intérêt à incorporer une compétence technologique dans une esthétique de l'authenticité qui la caractérise, et fait sa valeur. Le rattachement à une origine radiophonique semble nécessaire, afin de figurer une légitimité, tout en développant des contenus qui enrichissent ou complètent l'écoute du flux en direct, aux nombreux effets et qualités. Si la radio sait incarner une personnalité — la voix des animateurs d'une station fait écho aux intériorités des auditeurs —, ajouter une dimension technologique à ses compétences lui permettra non plus de subir mais d'utiliser les mécanismes de la personnalisation algorithmique à son avantage. En d'autres termes, il s'agit de faciliter l'accès de l'utilisateur à un média, puis à un contenu qui propose une expérience unique, tout en lui véhiculant l'authenticité d'une radio d'abord incarnée par ses voix.

La progressive artificialisation des pratiques médiatiques s'accompagne pour les entreprises du secteur d'une hybridation exponentielle de leurs métiers entre une mission d'information ou un désir de divertissement des publics et une compétence technologique. Peu développé dans le cas des enceintes connectées, ce phénomène est perceptible dans l'évolution du métier de journaliste de presse : auparavant restreint à son support d'expression, il est aujourd'hui capable de réaliser des vidéos, de l'audio, et d'incorporer des mécaniques de référencement web[271] ainsi que de partage sur les réseaux sociaux[272] à la construction même de son papier. Si la radio dispose d'un avantage grâce

[271] Le *Search Engine Optimising*, ou SEO, suppose notamment de répéter plusieurs fois certains mots-clés pour qu'un contenu apparaisse parmi les premiers résultats suggérés par un moteur de recherche.

[272] Nombre de rédactions requièrent de leur journalistes qu'en plus de leur article, ils proposent un titre et une phrase d'accroche à publier sur les réseaux sociaux.

à l'ancrage de son identité et de ses savoir-faire dans la voix, peut-être perdra-t-elle cette course à l'acquisition de compétences technologiques. Les concentrations d'acteurs médiatiques dans des groupes diversifiés fragilisent les entités qui demeurent solitaires, mais affirment par-là leur indépendance. Cet aspect est visible lorsque le groupe radio RTL fusionne avec le géant télévisuel M6. Les ressources du groupe, décuplées, facilitent certainement l'acquisition de nouvelles compétences. Nous pourrions également mentionner le *Washington Post*, quotidien américain racheté par Amazon en 2013[273]. Le titre de presse bénéficie de l'offre technologique de sa maison-mère, qui diversifie ses sources de revenus et lui assure une capacité à produire des contenus hautement qualitatifs sur les enceintes connectées. Fort de ses compétences, le journal vend depuis 2014 ses solutions logicielles de publication en ligne à d'autres médias, par le biais de la firme Arc Publishing[274]. Sur Amazon Echo, le quotidien bénéficie d'un solide socle de revenus et de compétences technologiques lui permettant d'effectuer des économies d'échelle en développant des contenus distribués sur plusieurs plateformes sous forme de podcasts, mais tout à fait dédiés aux enceintes connectées[275]. Son « Retropod », qui retrace en cinq minute une histoire singulière, et son « Daily 202 », dans lequel un expert de la rédaction analyse une actualité politique en une petite dizaine de minutes, semblent pensés pour ce support et sont proposés sous forme de *skills*. Ce dernier exemple illustre la forte concurrence qui s'organise face à la légitimité des stations de radio dans le domaine de la voix, supposant qu'elles agissent, souvent avec moins de ressources.

---

[273] *Le Monde*, « Le patron d'Amazon rachète le "Washington Post" », 05/08/2013. Disponible à l'adresse suivante : https://abonnes.lemonde.fr/actualite-medias/article/2013/08/05/le-patron-d-amazon-rachete-le-washington-post_3457822_3236.html?

[274] *Fast Company*, « The Washington Post Is A Software Company Now », 17/11/2017. Disponible à l'adresse suivante : https://www.fastcompany.com/40495770/the-washington-post-is-a-software-company-now

[275] *Washington Post*, Page de présentation de l'offre de podcasts. Disponible à l'adresse suivante : https://www.washingtonpost.com/podcasts/?utm_term=.829b26b952b4

Conclusion.

Notre recherche sur la place de la création radiophonique sur les enceintes connectées arrive à son terme. Nous avons entrepris, par le prisme des évolutions du matériau vocal au cours de son transport dans une multiplicité d'espaces, de définir les contours de la remédiation de la radio en contexte de virtualité numérique, que les enceintes connectées illustrent et concrétisent. Cette étude a ainsi supposé que nous analysions les propriétés de la voix, avant d'en observer les évolutions en radio puis sur les enceintes connectées, nécessitant que nous caractérisions les modalités de ces supports (les acteurs qui les composent, les économies qu'ils instaurent) avant d'y qualifier le statut de la voix, dans une perspective d'exploration d'un marché en construction.

Notre première partie s'est concentrée sur les propriétés naturelles du matériau vocal, qui constituent selon nous les fondements de l'effet d'authenticité de la radio, et sont déployées par la virtualité numérique dans un double mouvement d'immersion et d'interaction. La voix a en effet de spécifique qu'en projetant le corps dans le monde, elle réunit les pôles opposés de l'intériorité (émotionnelle, gestuelle et musicale) et de la société (régie par un langage qui permet de communiquer avec les autres). Son transport dans des ondes radiophoniques suppose par conséquent à la fois une compétence des voix de radio — une « homo vox » — et un contrat de communication tacite avec les auditeurs. Ce premier travail radiophonique de la voix est porté par son dispositif d'émission et d'écoute, historiquement ancré dans les stations de la bande FM. Effets d'intimité, de communauté, de réel et d'imaginaire se rassemblent dans la perception d'une authenticité à l'écoute des voix de radio, qui se réalise pleinement dans un dispositif acousmatique et un flux en direct. La capacité de la voix de radio à immerger un auditeur dans le contenu, et à interagir avec lui à deux degrés (l'un, passif, de projection dans l'imaginaire et l'intériorité, l'autre, actif, de participation à l'antenne et de fédération de communautés), provient autant de son référent sensible que de son organisation dans le dispositif radiophonique. Ces qualités, la radio les partage avec la virtualité numérique, qui consiste en une entreprise d'immersion et d'interaction avec des utilisateurs dans une réalité annexe, voire complémentaire dans le cas d'un Internet dit symbiotique. De ce partage provient une osmose entre la radio et les réseaux sociaux, qui réagissent instantanément et fédèrent des communautés virtuelles. La virtualité numérique, car axée principalement sur l'image, menace cependant le dispositif acousmatique de la radio et son effet d'authenticité. Ce média, peu viral, est contraint de se dénaturer pour coller aux canons de la viralité en se morcelant, ou en ajoutant de la vidéo à son expression. À l'inverse, la saturation visuelle qui

caractérise la virtualité numérique est un atout pour la radio, média du temps long, de l'introspection et de l'accompagnement soutenu par un contexte d'hyperesthésie où les consommateurs sont guidés par leurs sensations davantage que leur raison — justifiant d'un attrait pour le grain de la voix de radio lié à la sensualité de son expression.

À la lumière de ces considérations, notre seconde partie a entrepris de définir la remédiation de la voix sur les enceintes connectées. Cette analyse comprend trois dimensions : économique (pour visualiser les forces en présence dans ce marché), médiatique (pour définir les positionnements des entreprises constituant un champ concurrentiel élargi pour la radio dans le maniement de la voix), et technologique (pour comprendre les enjeux que revêt le déplacement de la source de la voix d'un corps humain à une Intelligence Artificielle). Sans surprise, le territoire où le marché des enceintes connectées est le plus développé est les États-Unis, berceau de cette technologie développée par des GAFAM. Ces derniers subissent cependant la concurrence d'autres groupes à l'échelle internationale, notamment asiatiques, supposant des stratégies de mutualisation des ressources telles que l'alliance entre Microsoft Cortana et Amazon Alexa. Le degré d'implantation de ces acteurs sur un territoire a une incidence directe sur la richesse — tant en termes de qualité que de quantité — de l'offre des médias sur les enceintes connectées. Si les rédactions de nombreux médias américains dédient des équipes au travail de la voix sur ces nouveaux supports, portées par un marché florissant, l'offre des médias en France demeure peu originale et puissante. Notre analyse des *skills* des médias associés au lancement d'Amazon Echo en France le 13 juin 2018, adjointe d'une étude de l'offre des stations de radio sur ce même support à l'été 2018, fait montre de ce faible développement. La voix apparaît comme l'opportunité d'une diversification de marques médias vécues comme globales, prenant la parole sur une multitude de supports numériques, mais n'investissant paradoxalement pas pleinement un marché encore embryonnaire dans l'hexagone. Les stations de radio, qui concentrent notre intérêt car leurs identités sont ancrées dans le travail de la voix, ne proposent pas de contenus profitant de l'interaction directe que les enceintes connectées permettent avec les utilisateurs. Nous en avons conclu que ce média est en cours de redéfinition, dans un souhait de trouver le juste équilibre entre le flux en direct — qui représente la majorité de ses audiences — et son offre numérique.

Parallèlement à cela, les enceintes connectées déplacent le matériau vocal d'un corps — le locuteur, radiophonique ou non — à une machine. Ce mouvement dépasse l'altération du matériau vocal effectuée par la radio que Pierre Schaeffer explique avec la notion de « dynéma », puisqu'il en modifie la source. Il a des répercussions audibles — en termes d'élocution et de compréhension, complexifiant la possibilité d'une conversation entre l'assistant vocal et son propriétaire — que les

concepteurs d'Intelligences Artificielles entendent pallier sans pour l'instant y parvenir. L'incorporation de travail humain dans des procédés de *machine learning*, censés attester de la perfection des enceintes connectées, souligne les progrès qui restent à accomplir. Si l'artificialité parvient aujourd'hui à répliquer l'intelligence, elle n'a pas conquis les émotions. Ces dernières, incarnées dans un corps, permettent la modulation de la voix — ces sonorités étant elles-mêmes porteuses de sens — et dotent le langage de sa capacité de connotation, trop complexe pour une Intelligence Artificielle. L'imperfection de la machine n'est néanmoins pas qu'involontaire. En effet, le principe de la *uncanny valley* (« la vallée dérangeante »), fondateur en robotique, stipule qu'une machine se rapprochant trop d'une ressemblance humaine sera monstrueuse dans ses défauts, révélations d'une tromperie, suscitant le rejet de l'utilisateur. Il s'agit par conséquent pour les concepteurs des intelligences qui se matérialisent dans les enceintes connectées d'en améliorer les facultés conversationnelles (d'élocution et de compréhension) pour fluidifier l'ancrage de ces assistants dans le quotidien, tout en signifiant subtilement qu'il s'agit de machines. Les sonorités robotiques de SIRI, assistant vocal disponible dans les appareils Apple, en sont un exemple.

La radio se situe par conséquent à un carrefour des transformations médiatiques, que l'altération de la voix sur les enceintes connectées cristallise — les pratiques médiatiques s'artificialisent, et une pluralité d'acteurs prétendent aux mêmes savoir-faire, leur faisant perdre en spécificité et nécessitant un important travail d'image. Notre troisième partie nous a permis de saisir cet enjeu — entre fonte dans un audio numérique et signification de l'attachement à un support —, et de le ramener à ce qui fait la différence et la valeur d'un média : ses savoir-faire. La radio diffère d'une marque commerciale en ce que ses prises de parole proviennent d'une multiplicité de corps, d'appréciations, et de conceptions du monde. C'est par conséquent la perception que les personnes travaillant en radio ont de leurs métiers qui définit les positionnements des marques qui composent ce média au sein de l'économie numérique. Nous avons pour cette raison décidé d'interroger une pluralité de profils au groupe radio RTL, avec une volontaire surreprésentation des travailleurs du son. Les notions de forme (pour désigner un matériau vocal fondateur d'une vision du monde axée autour du son) et de mise en forme (le fait d'opérer cette vision au sein d'une pratique professionnelle) sont au cœur de l'évolution de leurs métiers sur le numérique. Nos entretiens nous ont permis de constater que si la voix est perçue comme un fondement de la pragmatique radiophonique, sa redéfinition sur les enceintes connectées — et plus largement en contexte de virtualité numérique — n'est pas corrélée aux futurs de ces métiers par les personnes qui les exercent.

Cette synthèse de notre recherche nous permet d'aborder les hypothèses qui ont structuré notre problématique et nos réflexions.

Notre première hypothèse arguait que les technologies vocales font l'objet d'un engouement des entreprises de médias émanant d'une double opportunité reflétée dans les discours d'escorte des acteurs qui s'y engouffrent : la possibilité d'atteindre les publics dans de nouvelles situations d'écoute, et la diversification d'une offre éditoriale avec un *médium* aux atours plus authentiques dans une économie de l'attention caractérisée par un effet de saturation visuelle. Elle est confirmée par la qualification des propriétés du matériau vocal, au fondement de l'authenticité du média radio, qu'une pluralité de marques s'approprient sur les enceintes connectées. La voix permet à des marques qui en découvrent l'usage d'atteindre des utilisateurs dans de nouvelles situations d'écoutes — caractérisées par une fonction d'accompagnement, acousmatique —, et de déployer leur champ d'expression à de nouveaux canaux du numérique dans une logique de média global.

Notre seconde hypothèse souhaitait évaluer l'idée que sur les enceintes connectées, l'authenticité permise par la voix est altérée par deux éléments concomitants : l'interaction qui s'instaure entre l'utilisateur et une Intelligence Artificielle d'une part, et la création de contenus médiatiques adaptés à cette relation d'autre part. Elle semble elle aussi être confirmée, car l'effet d'authenticité que la voix permet en contexte radiophonique est altéré par l'usage d'une voix artificielle, qui encadre ou remplace la vive voix dans le cas du *text-to-speech.* Notre hypothèse doit cependant être nuancée car la voix est altérée dès son transport dans les ondes radiophoniques, comme la notion de « dynéma » l'illustre. L'instauration d'une distance entre son élocution et son écoute rompt un continuum spatio-temporel que tout le talent de la création radiophonique consiste à effacer, et que les concepteurs des enceintes connectées tentent de combler en figurant une personnalité artificielle qui répond à des logiques complexes, comme notre synthèse l'a rappelé.

Notre troisième hypothèse entendait qualifier le statut de la radio sur le numérique, entre remédiation et convergence, visualisée dans une position ambivalente dans le champ concurrentiel des enceintes connectées, supposant que sa puissance historique et économique est contrebalancée par la menace d'acteurs plus petits mais dédiés à la création digitale et par le déplacement de la voix dans une personnalité artificielle, allant jusqu'à redéfinir la pratique des métiers de radio. Cette hypothèse voit l'une de ses composantes invalidées. Si la radio se trouve dans une position ambivalente dans le contexte de l'économie de la voix numérique — dont les enceintes connectées constituent un aboutissement — qui suppose une redéfinition des métiers de radio, ces derniers ne semblent que peu affectés par ces changements. La progressive artificialisation des pratiques

médiatiques, si elle est effective, n'est pas complètement corrélée aux évolutions des métiers des personnes que nous avons interrogées.

L'engouement des entreprises médiatiques pour les outils mobilisant la voix, dont les enceintes connectées constituent un aboutissement en contexte numérique, construit-il des offres au sein desquelles les acteurs historiques de la création radiophonique se fondent dans le champ plus large de l'audio numérique ? La réponse est plurielle, et non définitive. Les enceintes connectées construisent un paysage médiatique où la radio perd en spécificité et est dénaturée par des voix factices. Les stations historiques s'y affaiblissent, concurrencées par d'autres médias et destituées de leur travail de la vive voix par des Intelligences Artificielles. La fonte de la radio dans l'audio numérique est prégnante à l'échelle du média — de la substance radiophonique —, qui prolifère dans des supports extérieurs au flux en direct des stations de la bande FM. La redéfinition des identités des stations de radio est une conséquence de ce contexte, mais leur force historique demeure prégnante face à des acteurs du numérique qui manquent soit de puissance (par exemple les studios de podcasts natifs), soit de compétences (l'exemple des titres de presse sur les enceintes connectées, qui n'affrontent pas la radio sur son terrain et prennent des voies détournées, l'illustre).

La remédiation de la substance radiophonique en une multiplicité de voix — qu'il s'agisse d'incarnations de médias globaux, des tessitures des Intelligences Artificielles, ou d'une radio dont les stations s'affaiblissent et perdent dans ce mouvement en spécificité — est en cours, et supposera probablement de s'y pencher à nouveau dans quelques années. La photographie du marché des enceintes connectées que nous avons entrepris de capturer en 2018 verra ses contours évoluer avec le temps. Elle nous a cependant permis d'aborder la place de la radio dans un environnement qui altère ses qualités sous une pluralité d'angles de vue, et pourrait supposer qu'elle acquière des compétences technologiques dans ce contexte de remédiation, comme nos recommandations à l'égard des stations le soulignent. Notre analyse témoigne en cela d'une transformation générale des entreprises médiatiques sur le numérique, aux pratiques accélérées, virtualisées et remodelées, pour s'adapter à une évolution des usages qui provient autant des avancées technologiques que de la perception des personnes qui construisent et consomment des médias. Plus largement, nous avons tenté de questionner la place du corps en contexte de virtualité numérique. Comment est-il dématérialisé ? Comment signifier sa présence ? À quel degré de désincarnation les publics sont-ils préparés ? Ce sont autant de questions abordées par le cas de la place de la radio sur les enceintes connectées, que notre recherche a tenté de qualifier tout en étant comprise dans un mouvement plus ample, dont elle n'effleure que la surface.

Bibliographie & Sitographie.

Ouvrages académiques

Antoine, F., et. alii, *Analyser la Radio. Méthodes et Mises en Pratique*, éd. De Boeck, 2016.

Bachelard, G., « Rêverie et Radio », in *Droit de Rêver*, éd. PUF, 1970.

Barthes, R., *Le Bruissement de la Langue*, éd. Seuil, 1984.

Barthes, R., *La Chambre Claire. Note sur la Photographie*, éd. Seuil, 1980.

Barthes, R., *Mythologies*, éd. Seuil, 1957.

Barthes, R., *Le Plaisir du Texte*, éd. Seuil, 1973.

Barthes, R., *Le Grain de la Voix. Entretiens 1962-1980*, éd. Seuil, 1981.

Baudrillard, J., *Le Système des Objets*, éd. Gallimard, 1968.

Bell, A., « Language Style as Audience Design », in *Language in Society*, 13(2), pp.145-204.

Benjamin, W., *L'Oeuvre d'Art à l'Époque de sa Reproductibilité Technique*, 1955, éd. Allia, 2016.

Boltanski, L., *La Souffrance à Distance. Morale Humaine, Médias et Politique*, éd. Métailié, 1993.

Bolter, J.D., Grusin, R., *Remediation. Understanding New Media*, MIT Press, 1998.

Candel, E., Jeanne-Perrier, V., Souchier, E., « Petite Formes, Grands Desseins. D'une Grammaire des Énoncés Éditoriaux à la Standardisation des Écritures », in Davallon, J., (dir.), *L'Économie des Écritures sur le Web*, éd. Hermès-Lavoisier, 2012, pp.165-201.

Castarède, M.-F., *La Voix et ses Sortilèges*, éd. Les Belles Lettres, 2004.

Charaudeau, P., (dir.), *Aspects du Discours Radiophonique*, éd. Didier Érudition, 1984.

Charles, D., *Le Temps de la Voix*, éd. Delarge, 1978.

Comte, A., *Discours sur l'Esprit Positif*, 1844, éd. Hachette Livre BNF, 2013.

Dayan, D., Katz, E., *La Télévision Cérémonielle. Anthropologie et Histoire en Direct*, éd. PUF, 1996.

Derrida, J., *De la Grammatologie*, éd. de Minuit, 1967.

Dyson, F., *Sounding New Media. Immersion and Embodiment in the Arts and Culture*, éd. University of California Press, 2009.

Fónagy, I., *La Vive Voix. Essais de Psycho-phonétique*, éd. Payot, 1983.
Gallagher, R., « Générer l'euphorie en ligne : l'esthétique de la culture ASMR », in *Audimat*, vol.9, 2018.

Genette, G., « Langage Poétique, Poétique du Langage », in *Figures II*, éd. Seuil, 1979, pp.123-153.

Glévarec, H., *France Culture à l'Oeuvre. Dynamique des Professions et Mise en Forme Radiophonique*, éd. CNRS, 2001.

Glévarec, H., « Le Propre de la Radio. Fonctions Radiophoniques et Nouveau Statut de la Radio dans l'Environnement Numérique », in *Le Temps des Médias*, n°22, 2014.

Goffman, E., *La Mise en Scène de la Vie Quotidienne. 1 : La Présentation de Soi*, éd. Minuit, 1973.

Goffman, E., « Radio Talk », in *Forms of Talk*, éd. University of Pennsylvania Press, 1981, pp. 197-327.

Howes, D., « Hyperaesthesia, or The Sensual Logic of Late Capitalism », in *Empire of the Senses*, éd. David Howes, 2004.

Jenkins, H., *Convergence Culture: Where Old and New Media Collide*, New York University Press, 2006.

Leroi-Gourhan, A., *Le Geste et la Parole*, éd. Albin Michel, tome 2, 1965.

Marcus, G., *Three Songs, Three Singers, Three Nations*, trad. de l'Américain par Godard, G., éd. Allia, 2018.

McLuhan, M., *Pour Comprendre les Médias*, éd. Seuil, 1968.

Metz, C., « Le Perçu et le Nommé », *Essais Sémiotiques*, éd. Klincksieck, 1977, pp. 129-163.

Paranthoën, Y., *Propos d'un Tailleur de Sons*, éd. Phonurgia Nova, 2002.

Piette, A., *Ethnographie de l'Action. L'Observation des Détails*, éd. Métailié, 1996.

Rheingold, H., *Les Communautés Virtuelles*, éd. Addison Wesley France, 1996.

Rousseau, J.-J., *Essai sur l'Origine des Langues*, éd. Aubier, 1973.

Storr, A., *Music and the Mind*, éd. Harper Collins, 1993.

Articles universitaires

Barthes, R., *Éléments de Sémiologie*, in *Communications*, 4, 1964, « Recherches Sémiologiques », pp.91-135. Disponible à l'adresse suivante : https://www.persee.fr/doc/comm_0588-8018_1964_num_4_1_1029

Boula de Mareüil, P., « Qu'est ce qu'un phonostyle ? », in *Nouveaux Cahiers de Linguistique Française*, 2014, 31, pp.9-19. Disponible à l'adresse suivante : http://clf.unige.ch/files/9014/4102/7363/1_Boula_9-19.pdf

De Iulio, S., « Lev Manovitch, *The Language of New Media* », in *Questions de Communication*, 4, 2003, pp.473-475. Disponible à l'adresse suivante : https://journals.openedition.org/questionsdecommunication/5949

INA Global, Santolaria, N., « Interfaces vocales : attention, dangers ! », 06/12/2017. Disponible à l'adresse suivante : https://www.inaglobal.fr/numerique/article/interfaces-vocales-attention-dangers-10013

INA Global, Smati, N., Fievet, A.-C., « A la radio, la voix donne à écouter et à voir », 04/12/2017. Disponible à l'adresse suivante : https://www.inaglobal.fr/radio/article/la-radio-la-voix-donne-ecouter-et-voir-10010

Le Tinnier, F., « Dynamique de la voix enregistrée selon Pierre Schaeffer (1941-1963) », in *Entrelacs*, 2014. Disponible à l'adresse suivante : https://journals.openedition.org/entrelacs/621

Patrin-Leclère, V., « Un média est-il une marque ? », in *Communication*, Vol. 32/2, 2013, mis en ligne le 07 avril 2014. Disponible à l'adresse suivante : https://journals.openedition.org/communication/5027

Rieusset-Lemarié, I., « Entre la « main de l'œil » et l' « œil digital », proximité et profondeur : la dimension haptique à l'horizon du Cinéma 3D et des acteurs virtuels », *Entrelacs*, 2013. Disponible à l'adresse suivante : https://journals.openedition.org/entrelacs/538

Tétu, J.F., « La radio, un média délaissé », *Hermès, La Revue*, 2004/1 (n° 38), p. 63-69. Disponible à l'adresse suivante : https://www.cairn.info/revue-hermes-la-revue-2004-1-page-63.htm

Articles de presse

20 Minutes, Beaudonnet, L., « VIDEO. Alexa en France : Retrouvez toute l'actu de « 20 Minutes » sur Alexa d'Amazon », 22/06/2018. Disponible à l'adresse suivante : https://www.20minutes.fr/high-tech/2289259-20180622-video-alexa-france-retrouvez-toute-actu-20-minutes-alexa-amazon

BFM Business, Henni, J., « Quelle est la radio dont les auditeurs sont les plus vieux? », 14/02/2016. Disponible à l'adresse suivante : https://bfmbusiness.bfmtv.com/entreprise/quelle-est-la-radio-dont-les-auditeurs-sont-les-plus-vieux-950088.html

Business Wire, « Juniper Research: Voice Assistants Used in Smart Homes to Grow 1,000%, Reaching 275 Million by 2023, as Alexa Leads the Way », 25/06/2018. Disponible à l'adresse suivante : https://www.businesswire.com/news/home/20180625005474/en/Juniper-Research-Voice-Assistants-Smart-Homes-Grow

CB News, Wojciak, T., « Amazon lance Alexa en France : les médias s'y engouffrent… », 06/06/2018. Disponible à l'adresse suivante : http://www.cbnews.fr/mobile/amazon-lance-alexa-en-france-les-medias-s-y-engouffrent-a1043088

CB News, Wojciak, T., « France Télévisions étend son partenariat avec Brut », 07/03/2017. Disponible à l'adresse suivante : http://www.cbnews.fr/digital/france-televisions-etend-son-partenariat-avec-brut-a1033931

CNet France, Zaffagni, M., « Amazon et Microsoft unissent leurs assistants Alexa et Cortana », 16/08/2018. Disponible à l'adresse suivante : https://www.cnetfrance.fr/news/amazon-et-microsoft-unissent-leurs-assistants-alexa-et-cortana-39872425.htm

Columbia Journalism Review, Brey, J., « Amid audio boom, magazines hire actors to voice their work », 05/01/2018. Disponible à l'adresse suivante : https://www.cjr.org/business_of_news/actors-magazines-voice.php

Digiday, Moses, L., « 'We're Pivoting to Words: *Slate* Says Podcasts Are Now 25 Percent of Revenue », 16/01/2018. Disponible à l'adresse suivante : https://digiday.com/media/slate-pivot-to-words/

Digiday, Moses, L., « For news publishers, smart speakers are the hot new platform », 23/08/2018. Disponible à l'adresse suivante : https://digiday.com/media/news-publishers-smart-speakers-hot-new-platform/

Fast Company, « The Washington Post Is A Software Company Now », 17/11/2017. Disponible à l'adresse suivante : https://www.fastcompany.com/40495770/the-washington-post-is-a-software-company-now

Franceinfo, « franceinfo, n°1 de l'info en continu sur les assistants vocaux, disponible via les services d'Amazon », 13/06/2018. Disponible à l'adresse suivante : https://www.francetvinfo.fr/internet/amazon/franceinfo-n1-de-l-info-en-continu-sur-les-assistants-vocaux-disponible-via-les-services-d-amazon_2798817.html

iGeneration, Moussie, S., « N'achetez pas le Google Home Mini… il est offert partout », 07/12/2017. Disponible à l'adresse suivante : https://www.igen.fr/accessoires/2017/12/nachetez-pas-le-google-home-mini-il-est-offert-partout-102247

Journal du Net, « Echo / Alexa : l'enceinte d'Amazon à la loupe [prix, apps, config…] », 23/08/2018. Disponible à l'adresse suivante : https://www.journaldunet.com/ebusiness/internet-mobile/1194132-echo-alexa-l-enceinte-d-amazon-a-la-loupe-prix-apps-config/

L'ADN, Huot, A., « Lancement d'Alexa en France : quels médias ont répondu présent ? », 13/06/2018. Disponible à l'adresse suivante : http://www.ladn.eu/media-mutants/les-medias-francais-sur-alexa/

L'ADN, Robillart, O., « On n'aura bientôt plus besoin d'écrire... La révolution de la voix est en marche », 02/02/2018. Disponible à l'adresse suivante : http://www.ladn.eu/tech-a-suivre/ia-et-chatbot/reconnaissance-vocale-les-start-up-en-france/

LCI, Ingrand, C., « 'Alexa, quelles sont les nouvelles ?' : LCI vous informe sur l'assistant vocal d'Amazon », 14/06/2018. Disponible à l'adresse suivante : https://www.lci.fr/high-tech/alexa-quelles-sont-les-nouvelles-lci-vous-accompagne-et-vous-informe-sur-l-assistant-vocal-d-amazon-2090470.html

Les Échos, Alcaraz, M., « Les médias parient sur les enceintes connectées », 13/06/2018, n°22715.

Le Figaro, « Mathieu Gallet : « La prochaine révolution numérique sera la voix », 25/04/2017. Disponible à l'adresse suivante : http://www.lefigaro.fr/medias/2017/04/23/20004-20170423ARTFIG00111-mathieu-galletla-prochaine-revolution-numerique-sera-la-voix.php

Le Huff Post, Phalippou, A., « Le HuffPost se lance dans l'audio sur Amazon Echo », 12/06/2018. Disponible à l'adresse suivante : https://www.huffingtonpost.fr/alexandre-phalippou/le-huffpost-se-lance-dans-l-audio-sur-amazon-echo_a_23450332/

Le Monde, « Le patron d'Amazon rachète le "Washington Post" », 05/08/2013. Disponible à l'adresse suivante : https://abonnes.lemonde.fr/actualite-medias/article/2013/08/05/le-patron-d-amazon-rachete-le-washington-post_3457822_3236.html?

Le Monde, « M6 va racheter les radios de RTL pour 216 millions d'euros », 14/12/2016. Disponible à l'adresse suivante : https://www.lemonde.fr/actualite-medias/article/2016/12/13/m6-ouvre-des-negociations-pour-racheter-rtl-group-en-france-pour-216-millions-d-euros_5048419_3236.html

Le Monde, Six, N., « Comparatif : HomePod, Echo, Home… Quelle enceinte connectée choisir ? », 23/06/2018. Disponible à l'adresse suivante : https://www.lemonde.fr/pixels/article/2018/06/23/comparatif-homepod-echo-home-quelle-enceinte-connectee-choisir_5320100_4408996.html

Les Numériques, « Samsung Galaxy Home, l'enceinte connectée rivale du HomePod », 10/08/2018. Disponible à l'adresse suivante : https://www.lesnumeriques.com/assistant-domotique/samsung-galaxy-home-p45463/samsung-devoile-galaxy-home-enceinte-connectee-rivale-homepod-n77029.html

Méta-Média, Chevet, C., « Dialoguiste pour robot : l'humain derrière la machine », 06/06/2018. Disponible à l'adresse suivante : https://www.meta-media.fr/2018/06/06/dialoguiste-pour-robot-lhumain-derriere-la-machine.html

New York Magazine - Daily Intelligencer, « What's Behind the Great Podcast Renaissance? », 31/10/2014. Disponible à l'adresse suivante : http://nymag.com/daily/intelligencer/2014/10/whats-behind-the-great-podcast-renaissance.html

Nieman Lab, Barot, T., « The future of news is humans talking to machines », 18/09/2018. Disponible à l'adresse suivante : http://www.niemanlab.org/2017/09/the-future-of-news-is-humans-talking-to-machines/

Numerama, Turcan, M., « Comment Amazon Echo s'appuie sur les médias pour conquérir les Français », 06/06/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/380039-comment-amazon-echo-sappuie-sur-les-medias-pour-conquerir-les-francais.html
Numerama, Turcan, M., « Comment les assistants vocaux vont devenir nos « compagnons de vie » », 26/05/2018. Disponible à l'adresse suivante : https://www.numerama.com/tech/378318-comment-les-assistants-vocaux-vont-devenir-nos-compagnons-de-vie.html

Offre Média, « Le potentiel du marché des assistants vocaux étudié par Mindshare », 27/04/2018. Disponible à l'adresse suivante : https://www.offremedia.com/le-potentiel-du-marche-des-assistants-vocaux-etudie-par-mindshare

Offre Média, « Les supports multimédias atteignent 12,2% du volume total d'écoute de la radio selon Médiamétrie », 18/05/2018. Disponible à l'adresse suivante : https://www.offremedia.com/les-supports-multimedias-atteignent-122-du-volume-total-decoute-de-la-radio-selon-mediametrie

Oregon Live, Acker, L., « Portland woman says Amazon Alexa recorded personal conversation, sent it to contact », 25/05/2018. Disponible à l'adresse suivante : https://www.oregonlive.com/trending/2018/05/portland_woman_says_amazon_ale.html

PR Newswire, « Voice Shopping Set to Jump to $40 Billion By 2022, Rising From $2 Billion Today », 28/02/2018. Disponible à l'adresse suivante : https://www.prnewswire.com/news-releases/voice-shopping-set-to-jump-to-40-billion-by-2022-rising-from-2-billion-today-300605596.html

Quartz, Gerhart, A., « Alexa makes decision-making easier than ever—by making your choices for you », 10/07/2018. Disponible à l'adresse suivante : https://qz.com/1323884/alexa-makes-decision-making-easier-than-ever-by-making-your-choices-for-you/

Slate, Kroh, L.,« Pourquoi est ce que la radio n'est pas virale ? », 28/02/2014. Disponible à l'adresse suivante : http://www.slate.fr/story/84009/radio-virale

Stratégies, « 6% des Français ont une enceinte connectée », 23/05/2018. Disponible à l'adresse suivante : http://www.strategies.fr/actualites/medias/4013057W/6-des-francais-ont-une-enceinte-connectee.html

Stratégies, Fraioli, B., « L'ad-switching à l'ère du médiaplanning », 20/04/2017. Disponible à l'adresse suivante : http://www.strategies.fr/actualites/medias/1062735W/l-ad-switching-a-l-ere-du-mediaplanning.html

Syntone, Noiseau, E., « Vers un art radiophonique numérique ? », 18/04/2013. Disponible à l'adresse suivante : http://syntone.fr/vers-un-art-radiophonique-numerique-4-4/

Ted Crunch, Constine, J., « Micropodcasting? Facebook tries Voice Clip status updates », 02/03/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/03/02/facebook-voice-clips/

Ted Crunch, Ha, A., « Bloomberg Media Group's Chief Product Officer sees big opportunities in audio », 07/08/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/06/07/bloomberg-doubles-down-on-audio/?guccounter=1

Ted Crunch, Koponen, J., « A new hope: AI for news media », 12/07/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/07/12/a-new-hope-ai-for-news-media/

Ted Crunch, Perez, S., « Amazon launches a Polly WordPress plugin that turns blog posts into audio, including podcasts », 08/02/2018. Disponible à l'adresse suivante : https://techcrunch.com/2018/02/08/amazon-launches-a-polly-wordpress-plugin-that-turns-blog-posts-into-audio-including-podcasts/

Ted Crunch, Perez, S., « Spoken Editions goes live on iTunes so you can listen to your favorite websites », 03/10/2016. Disponible à l'adresse suivante : https://techcrunch.com/2016/10/03/spoken-editions-go-live-on-itunes-so-you-can-listen-to-your-favorite-websites/

The Guardian, Lay, S., « Uncanny valley: why we find human-like robots and dolls so creepy », 13/11/2015. Disponible à l'adresse suivante : https://www.theguardian.com/commentisfree/2015/nov/13/robots-human-uncanny-valley

The Guardian, Solon, O., « The rise of 'pseudo-AI': how tech firms quietly use humans to do bots' work », 06/07/2018. Disponible à l'adresse suivante : https://www.theguardian.com/technology/2018/jul/06/artificial-intelligence-ai-humans-bots-tech-companies

The Inquirer, Page, C., « Facebook 'delays smart speaker launch' in wake of Cambridge Analytica scandal », 24/04/2018. Disponible à l'adresse suivante : https://www.theinquirer.net/inquirer/news/3030793/facebook-delays-smart-speaker-launch-in-wake-of-cambridge-analytica-scandal

The New York Times, Chokshi, N., « Amazon's Audible Goes Long on Short-Form Audio », 07/07/2016. Disponible à l'adresse suivante : https://www.nytimes.com/2016/07/08/business/media/amazons-audible-goes-long-on-short-term-audio.html

The New York Times, « Welcome to the Post-Text Future », 14/02/2018. Disponible à l'adresse suivante : https://www.nytimes.com/interactive/2018/02/09/technology/the-rise-of-a-visual-internet.html

The Verge, Vincent, J., « Inside Amazon's $3.5 million competition to make Alexa chat like a human », 13/06/2018. Disponible à l'adresse suivante : https://www.theverge.com/2018/6/13/17453994/amazon-alexa-prize-2018-competition-conversational-ai-chatbots

Voicebot, Kinsella, B., « Amazon Echo and Google Home European Smart Speaker Sales Approach 6.5 Million Units in 2017 », 28/02/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/02/28/amazon-echo-google-home-european-smart-speaker-sales-approach-6-5-million-units-2017/

Voicebot, Kinsella, B., « Amazon Closes Year With 266% Alexa Skill Growth in U.S. », 08/01/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/01/08/amazon-closes-year-with-266-alexa-skill-growth-u-s/

Voicebot, Kinsella, B., « U.S. Falls to 46% of Smart Speaker Sales in Q1 2018 on the Rise of Alibaba, Xiaomi and Korean Device Makers », 23/05/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/05/23/u-s-falls-to-46-of-smart-speaker-sales-in-q1-2018-on-the-rise-of-alibaba-xiaomi-and-korean-device-makers/

Voicebot, Kinsella, B., « US Smart Speaker Market Share : Apple Debuts at 4.1%, Amazon Falls 10 Points and Google Rises », 03/06/2018. Disponible à l'adresse suivante : https://voicebot.ai/2018/06/03/u-s-smart-speaker-market-share-apple-debuts-at-4-1-amazon-falls-10-points-and-google-rises/

Wired, Barrett, B., « Inside Amazon's Painstaking Pursuit to Teach Alexa French », 13/06/2018. Disponible à l'adresse suivante : https://www.wired.com/story/how-amazon-taught-alexa-to-speak-french/?mbid=nl_061318_daily_list_p

ZD Net, Aubert, P., « Google Home passe devant Amazon Echo », 29/05/2018. Disponible à l'adresse suivante : https://www.zdnet.fr/blogs/watch-it/google-home-passe-devant-amazon-echo-39868815.htm

Communiqués de presse

Les Échos-Le Parisien, « Le groupe Les Échos-Le Parisien fait entendre sa voix sur Alexa, l'assistant vocal d'Amazon », 06/06/2018. Disponible à l'adresse suivante : http://pressroom-lesechos-leparisien.com/2018/06/06/le-groupe-les-echos-le-parisien-fait-entendre-sa-voix-sur-alexa-lassistant-vocal-damazon/

L'Équipe, « L'Équipe lance son flash briefing sur Amazon Echo », 06/06/2018. Disponible à l'adresse suivante : https://blog.lequipe.fr/operations/lequipe-amazon-echo/

Radio France, « Radio France poursuit son développement sur les assistants vocaux avec Amazon Alexa », 13/06/2018. Disponible à l'adresse suivante : http://www.radiofrance.fr/espace-pro/espace-presse/communiques-de-presse/2018/06/radio-france-poursuit-son-developpement-sur

Vidéos

Culture Pub, « Apple : Welcome Home - États-Unis 2018 ». Disponible à l'adresse suivante : http://www.culturepub.fr/videos/apple-homepod-welcome-home/

Culture Pub, « Google Home - États-Unis, 2017 ». Disponible à l'adresse suivante : « http://www.culturepub.fr/videos/google-google-home-google-home/

YouTube, compte officiel Amazon, « Alexa Loses Her Voice - Amazon Super Bowl LII Commercial », Mis en ligne le 31/01/2018. Disponible à l'adresse suivante : https://youtu.be/J6-8DQALGt4

YouTube, « Geeshie Wiley - Last Kind Word Blues ». Mis en ligne le 24/05/2009. Disponible à l'adresse suivante : https://youtu.be/oAKfy2W70Qg

Études

Edison Research, « The Infinite Dial 2018 », 08/03/2018. Disponible à l'adresse suivante : http://www.edisonresearch.com/infinite-dial-2018/

Edison Research, « The Podcast Consumer 2018 », 19/04/2018. Disponible à l'adresse suivante : http://www.edisonresearch.com/podcast-consumer-2018/

Edison Research & National Public Radio, « The Smart Audio Report - Spring 2018 ». Disponible à l'adresse suivante : https://www.nationalpublicmedia.com/smart-audio-report/latest-report/

Kantar France Insights, « Baromètre 2018 de la confiance des Français dans les Médias », 23/01/2018. Disponible à l'adresse suivante : https://fr.kantar.com/médias/digital/2018/barometre-2018-de-la-confiance-des-francais-dans-les-media/

Ministère de la Culture, « Médias et Publicité en Ligne », 25/07/2018. Disponible à l'adresse suivante : http://www.culture.gouv.fr/Thematiques/Audiovisuel/Rapports-etudes/Etude-sur-les-Medias-et-la-publicite-en-ligne-transfert-de-valeur-et-nouvelles-pratiques2

Statista, Boittiaux, P., « Apple à la traîne sur le marché des enceintes connectées », 19/06/2018. Disponible à l'adresse suivante : https://fr.statista.com/infographie/12350/apple-a-la-traine-sur-le-marche-des-enceintes-connectees/

Autres

Amazon Alexa Prize - Page de Présentation. Disponible à l'adresse suivante : https://developer.amazon.com/alexaprize

Amazon The Mechanical Turk - Page de Présentation. Disponible à l'adresse suivante : https://www.mturk.com

Bloomberg, Rosenberg, D.S., « Foundations of Machine Learning ». Disponible à l'adresse suivante : https://bloomberg.github.io/foml/#home

France Culture - « Hasta Dente ! » - Page de Présentation. Disponible à l'adresse suivante : https://www.franceculture.fr/evenement/hasta-dente

Glossaire E-Marketing, définition de « marque ». Disponible à l'adresse suivante : http://www.e-marketing.fr/Definitions-Glossaire/Marque-238333.htm

Le Big Data, L., B., « Machine Learning et Big Data : Définition et Explications », 06/07/2018. Disponible à l'adresse suivante : https://www.lebigdata.fr/machine-learning-et-big-data

Mougin, M., Mémoire de Master 2 Médias & Management, CELSA Sorbonne Université, « L'Intentionnalité narrative dans les émissions de radio », 2017.

NRJ, « NRJ sur Amazon Alexa ». Disponible à l'adresse suivante : http://www.nrj.fr/amazon-alexa

Radio France, Masterclass, « Smartspeakers / Assistants Vocaux : Comment se Faire Entendre ? », 11/06/2018. Disponible à l'adresse suivante : http://hyperradio.radiofrance.fr/innovation/inscrivez-a-masterclass-radio-france-station-f-8-juin/

Trésor de la Langue Française Informatisé. Disponible à l'adresse suivante : http://www.cnrtl.fr/definition/

Terra Nova, Gallet, M., « Transformation numérique, synergies agiles, responsabilité démocratique : trois enjeux pour les médias publics », 05/11/2017. Disponible à l'adresse suivante : http://tnova.fr/notes/transformation-numerique-synergies-agiles-responsabilite-democratique-trois-enjeux-pour-les-medias-publics

The New York Times - Podcasts : https://www.nytimes.com/podcasts/
WordPress Amazon Polly - Page de Présentation. Disponible à l'adresse suivante : https://wordpress.org/plugins/amazon-polly/

Washington Post, Page de présentation de l'offre de podcasts. Disponible à l'adresse suivante : https://www.washingtonpost.com/podcasts/?utm_term=.829b26b952b4

Annexes.

Annexe 1 — Guide d'entretien de Julien Le Ny et d'Alban Rivoal, producteurs à la station Fun Radio.

THEME 1 **PRESENTATION**	**Carrière**	**1.** Pourriez-vous vous présenter (métier, parcours) ? **2.** Quelles sont vos missions principales ? **3.** A quelles étapes du processus de mise en ondes intervenez-vous ?
	Situation professionnelle	**4.** Comment fonctionne le travail en équipe lors de la création radiophonique ? Les rôles sont-ils clairement distribués ? **5.** Certains aspects de votre métier changent-ils avec les évolutions des modes de consommation de la radio ?
THEME 2 **LE TRAVAIL DU SON ET DE LA VOIX EN CONTEXTE RADIOPHONIQUE**	**Travail du son**	**6.** Yann Paranthoën, opérateur du son et auteur de documentaires radiophoniques ayant exercé à Radio France, s'est décrit à plusieurs reprises comme un "tailleur de sons". Ce terme vous convient-il pour désigner votre métier, et pourquoi ? **7.** Comment parvient-on à retranscrire la "couleur" d'une station dans une production sonore ?
	Matérialité de la voix en radio	**8.** Quels sont selon vous le rôle et les effets de la voix en radio ? **9.** Comment qualifieriez-vous le rapport que l'écoute d'une voix instaure entre un animateur et ses auditeurs ? **10.** De quelle manière la voix intègre-t-elle votre réflexion dans les différentes étapes composant le travail de mise en ondes ? **11.** La voix est une émanation naturelle du corps. Sa modification et son travail en radio en altèrent-is certaines qualités, notamment émotionnelles, selon vous ? **12.** Comment conserve-t-on ce que des chercheurs ont appelé "l'effet d'humanité" d'une voix de radio dans des formats parfois très séquencés (type *playlist* pour les radios musicales) ?

THEME 3 **EVOLUTION DU METIER**	**La production de contenus délinéarisés**	**13.** Dans quelle mesure la désormais nécessaire production de contenus dédiés à une consommation délinéarisée, type podcasts, modifie-t-elle votre métier ? **14.** Incorporez-vous cette dimension (linéaire, non linéaire) dans votre processus de création ?
	Le travail de la voix et du son sur les enceintes connectés	**15.** Que pensez-vous du récent engouement autour des enceintes connectées en termes de productions médiatiques ? **16.** La radio a-t-elle un/des avantage(s) vis-à-vis d'autres médias (presse, TV…) sur les enceintes connectées ? Quel est/sont-il(s) ? **17.** La production de contenus dédiés aux enceintes connectées pourrait-elle modifier votre situation de travail (son fonctionnement, ses temporalités) ? **18.** Sur les enceintes connectées, une voix artificielle encadre l'expression radiophonique, ou bien s'y substitue dans le cas des articles écrits lus à l'oral, prisés des titres des presse sur ce support. Le mode d'expression radiophonique, axé sur la voix humaine, risque-t-il de perdre en authenticité ?
CONCLUSION	**Projection**	**19.** Dans un article publié dans *Les Échos* ("Les médias parient sur les enceintes connectées", 13 juin 2018 - n°2215), Charles-Emmanuel Bon, Directeur du Développement du pôle radio RTL, fait la déclaration suivante : "Les enceintes connectées pourraient devenir le relais de la radio dans le futur". Que veut-il dire selon vous ? **20.** Avec les enceintes connectées, la machine va jusqu'à remplacer la voix humaine (non sans problèmes techniques pour l'instant). Qu'est ce que cela présage des futures évolutions de votre métier ?

Annexe 2 — Retranscription de l'entretien mené avec Julien Le Ny, producteur à la station Fun Radio.

Annexe 3 — Retranscription de l'entretien mené avec Alban Rivoal, producteur à la station Fun Radio.

Les pages 148 à 158 ont été retirées de la version diffusée en ligne.

Annexe 4 — Guide d'entretien de Vladimir Bobrow et Sacha Ephraim, régisseurs-monteurs à la station RTL.

THEME 1 **PRESENTATION**	**Carrière**	**1.** Pourriez-vous vous présenter (métier, parcours) ? **2.** Quelles sont vos missions principales ? **3.** A quelles étapes du processus de mise en ondes intervenez-vous ?
	Situation professionnelle	**4.** Comment fonctionne le travail en équipe lors de la création radiophonique ? Les rôles sont-ils clairement distribués ? **5.** Certains aspects de votre métier changent-ils avec les évolutions des modes de consommation de la radio ?
THEME 2 **LE TRAVAIL DU SON ET DE LA VOIX EN CONTEXTE RADIOPHONIQUE**	**Travail du son**	**6.** Yann Paranthoën, opérateur du son et auteur de documentaires radiophoniques ayant exercé à Radio France, s'est décrit à plusieurs reprises comme un "tailleur de sons". Ce terme vous convient-il pour désigner votre métier, et pourquoi ? **7.** Comment parvient-on à retranscrire la "couleur" d'une station dans une production sonore ?
	Matérialité de la voix en radio	**8.** Quels sont selon vous le rôle et les effets de la voix en radio ? **9.** Comment qualifieriez-vous le rapport que l'écoute d'une voix instaure entre un animateur et ses auditeurs ? **10.** De quelle manière la voix intègre-t-elle votre réflexion dans les différentes étapes composant le travail de mise en ondes ? **11.** La voix est une émanation naturelle du corps. Sa modification et son travail en radio en altèrent-is certaines qualités, notamment émotionnelles, selon vous ? **12.** Comment conserve-t-on ce que des chercheurs ont appelé "l'effet d'humanité" d'une voix de radio dans des formats parfois très séquencés (type *playlist* pour les radios musicales) ?

THEME 3 **EVOLUTION DU METIER**	**La production de contenus délinéarisés**	**13.** Dans quelle mesure la désormais nécessaire production de contenus dédiés à une consommation délinéarisée, type podcasts, modifie-t-elle votre métier ? **14.** Incorporez-vous cette dimension (linéaire, non linéaire) dans votre processus de création ?
	Le travail de la voix et du son sur les enceintes connectés	**15.** Que pensez-vous du récent engouement autour des enceintes connectées en termes de productions médiatiques ? **16.** La radio a-t-elle un/des avantage(s) vis-à-vis d'autres médias (presse, TV…) sur les enceintes connectées ? Quel est/sont-il(s) ? **17.** La production de contenus dédiés aux enceintes connectées pourrait-elle modifier votre situation de travail (son fonctionnement, ses temporalités) ? **18.** Sur les enceintes connectées, une voix artificielle encadre l'expression radiophonique, ou bien s'y substitue dans le cas des articles écrits lus à l'oral, prisés des titres des presse sur ce support. Le mode d'expression radiophonique, axé sur la voix humaine, risque-t-il de perdre en authenticité ?
CONCLUSION	**Projection**	**19.** Dans un article publié dans *Les Échos* ("Les médias parient sur les enceintes connectées", 13 juin 2018 - n°2215), Charles-Emmanuel Bon, Directeur du Développement du pôle radio RTL, fait la déclaration suivante : "Les enceintes connectées pourraient devenir le relais de la radio dans le futur". Que veut-il dire selon vous ? **20.** Avec les enceintes connectées, la machine va jusqu'à remplacer la voix humaine (non sans problèmes techniques pour l'instant). Qu'est ce que cela présage des futures évolutions de votre métier ?

Annexe 5 — Retranscription de l'entretien mené avec Vladimir Bobrow, régisseur-monteur à la station RTL.

Annexe 6 — Retranscription de l'entretien mené avec Sacha Ephraim, régisseur-monteur à la station RTL.

Les pages 161 à 167 ont été retirées de la version diffusée en ligne.

Annexe 7 — Guide d'entretien de Rafaële Dhuicque-Filloux, directrice de la société de production Zéro Décibel.

THEME 1 **PRESENTATION**	**Carrière**	**1.** Pourriez-vous vous présenter (métier, parcours) ? **2.** Quelles sont vos missions principales ?
	Situation professionnelle	**3.** Comment fonctionne le travail en équipe lors de la création de spots publicitaires ? Les rôles sont-ils clairement distribués ? **4.** Certains aspects de votre métier changent-ils avec les évolutions des modes de consommation de la radio ?
THEME 2 **LE TRAVAIL DU SON ET DE LA VOIX EN CONTEXTE RADIOPHONIQUE**	**Travail du son**	**5.** Yann Paranthoën, opérateur du son et auteur de documentaires radiophoniques ayant exercé à Radio France, s'est décrit à plusieurs reprises comme un "tailleur de sons". Ce terme vous convient-il pour désigner votre métier, et pourquoi ? **6.** Essayez-vous de retranscrire ce que l'on nomme la "couleur" d'une station dans vos productions ?
	Matérialité de la voix en radio	**7.** Quels sont selon vous le rôle et les effets de la voix en radio et le rapport qui s'instaure entre un animateur et ses auditeurs ? **8.** De quelle manière la voix intègre-t-elle votre réflexion dans les différentes étapes composant le travail de mise en ondes, dont la finalité est l'efficacité commerciale ? **9.** La voix est une émanation naturelle du corps. Sa modification et son travail en radio en altèrent-is certaines qualités, notamment émotionnelles, selon vous ? **10.** Comment conserve-t-on ce que des chercheurs ont appelé "l'effet d'humanité" d'une voix de radio dans des formats très courts et séquencés ?

THEME 3 **EVOLUTION DU METIER**	**La production de contenus délinéarisés**	**13.** Dans quelle mesure la désormais nécessaire production de contenus dédiés à une consommation délinéarisée, type podcasts, modifie-t-elle votre métier ? **14.** Incorporez-vous cette dimension (linéaire, non linéaire) dans votre processus de création ?
	Le travail de la voix et du son sur les enceintes connectés	**15.** Que pensez-vous du récent engouement autour des enceintes connectées en termes de productions médiatiques ? **16.** La radio a-t-elle un/des avantage(s) vis-à-vis d'autres médias (presse, TV…) sur les enceintes connectées ? Quel est/sont-il(s) ? **17.** La production de contenus dédiés aux enceintes connectées pourrait-elle modifier votre situation de travail (son fonctionnement, ses temporalités) ? **18.** Sur les enceintes connectées, une voix artificielle encadre l'expression radiophonique, ou bien s'y substitue dans le cas des articles écrits lus à l'oral, prisés des titres des presse sur ce support. Le mode d'expression radiophonique, axé sur la voix humaine, risque-t-il de perdre en authenticité (et, dans un contexte publicitaire, en efficacité) ?
CONCLUSION	**Projection**	**19.** Qu'est ce que ce type d'innovations, qui renouvelle le rapport au média audio, pourrait présager des évolutions médiatiques et de la création publicitaire ?

Annexe 8 — Retranscription de l'entretien de Rafaële Dhuicque-Filloux, directrice de la société de production Zéro Décibel.

Les pages 170 à 175 ont été retirées de la version diffusée en ligne.

Annexe 9 — Guide d'entretien de Nicolas Dot, chef de produit marketing radio au groupe RTL, chargé d'une veille sur les enceintes connectées.

THÈME 1 PRÉSENTATION	**Profession**	**1.** Pourriez-vous vous présenter (profession, parcours) ? **2.** Quelles sont vos missions principales ? **3.** Avez-vous un goût particulier pour la radio ? Pour quelles raisons ? **4.** Dans quelle perspective avez-vous intégré la régie publicitaire plutôt que l'antenne ? Par choix, par opportunité ?
THÈME 2 LA VEILLE SUR LES ENCEINTES CONNECTÉES	**Principes et cadre de la veille**	**5.** Vous effectuez dans le cadre de vos missions une veille sur les assistants connectés. Pouvez-vous m'en expliquer le principe ? **6.** Quelle est la finalité de cette veille ?
	Technologie	**7.** Quel est le fonctionnement général d'une enceinte connectée, et quelles grandes différences observe-t-on entre les produits des trois leaders du secteur (Amazon, Google et Apple) ? **8.** Les fonctions les plus utilisées des enceintes connectées sont les questions d'ordre général sur internet, le minuteur et la musique. Pouvez-vous m'expliquer l'intérêt pour un média d'investir le champ des enceintes connectées ? **9.** Y a-t-il une innovation ou un fait qui vous a particulièrement marqué à propos des assistants vocaux ? **10.** En mai dernier, un couple habitant à Portland a découvert que leur enceinte connectée Amazon avait enregistré puis envoyé par erreur l'une de leurs conversations. Quel risque cela représente-il pour nos vies privées ? **11.** Sur les enceintes connectées, le contenu semble aussi important que la manière dont le consommateur peut y parvenir. Quelle est l'avancée des médias dans ce domaine ? **12.** Que pensez-vous du mode d'interaction permis par les enceintes connectées ? Y a-t-il un nouveau champ de création médiatique qui se déploie ?

THÈME 3 **ANALYSE STRATÉGIQUE**	**Ensemble des médias présents sur les enceintes connectées**	**13.** Qu'observez-vous à votre niveau quant au positionnement des médias sur les enceintes connectées : leurs discours marketing, la manière dont elles légitiment leur présence ? **14.** En termes de monétisation, la publicité classique, le sponsoring et les contenus de marque cohabitent sur ce support, très empreint des codes de l'audio numérique (notamment des podcasts). Quelle vous semble être l'option la plus empruntée, et laquelle vous semble la plus intéressante ?
	Place de la radio d'un point de vue éditorial	**15.** Pensez-vous que la radio a une légitimité particulière sur les enceintes connectées, issue de son identité de média ? **16.** Est-il possible de retranscrire ce que l'on nomme communément la « couleur » d'une station sur ces contenus courts, dénués du moindre support visuel (type logo) ? **17.** Les larges audiences de la radio, structurellement vieillissantes, lui confèrent un poids historique et économique. Sont-elles cependant une contrainte sur un support où le dynamisme et l'innovation semblent être de rigueur ?
CONCLUSION	**Projection**	**18.** Dans un article publié dans *Les Échos* (« Les médias parient sur les enceintes connectées », 13 juin 2018 – n°2215), Charles-Emmanuel Bon, Directeur du Développement du pôle radio RTL, fait la déclaration suivante : « Les enceintes connectées pourraient devenir le relais de la radio dans le futur ». Que veut-il dire selon vous ? **19.** Avec les enceintes connectées, la machine va jusqu'à remplacer la voix humaine (non sans problèmes techniques pour l'instant). Qu'est ce que cela présage des futures évolutions des contenus et des usages médiatiques ?

Annexe 10 — Retranscription de l'entretien de Nicolas Dot, chef de produit marketing radio au groupe RTL, chargé d'une veille sur les enceintes connectées.

Les pages 178 à 184 ont été retirées de la version diffusée en ligne.

Annexe 11 — Guide d'entretien de Pierre Laurent, chef de produit marketing radio au groupe RTL, en charge du Laboratoire Sonore.

THEME 1 **PRESENTATION**	**Carrière**	**1.** Pourriez-vous vous présenter (métier, parcours) ? **2.** Quelles sont vos missions principales ?
	Laboratoire Sonore	**3.** Vous êtes en charge du Laboratoire Sonore, qui établit des critères d'efficacité des messages publicitaires. Quel est son fonctionnement ?
THEME 2 **LES EFFETS DE LA RADIO**	**Pouvoir du son**	**4.** Quels sont selon vous les pouvoir de la voix et son, matériaux privilégiés de la radio ? **5.** Comment se traduit la notion de "couleur d'antenne" au sein de ce dispositif ?
	Matérialité de la voix en radio	**6.** Quels sont selon vous le rôle et les effets de la voix en radio et le rapport qui s'instaure entre un animateur et ses auditeurs ? **7.** De quelle manière la voix intègre-t-elle votre réflexion dans les différentes étapes composant le travail de mise en ondes, dont la finalité est l'efficacité commerciale ? **8.** La voix est une émanation naturelle du corps. Sa modification et son travail en radio en altèrent-is certaines qualités, notamment émotionnelles, selon vous ?
THEME 3 **ÉVOLUTION DU MÉDIA RADIO**	**La place de la radio dans l'environnement numérique**	**9.** Percevez-vous une redéfinition de la place de la radio sur le numérique ? **10.** Les enceintes connectées sont-elles exemplaires d'un affaiblissement des stations historiques ?
	Le travail de la voix et du son sur les enceintes connectés	**11.** Que pensez-vous du récent engouement autour des enceintes connectées en termes de productions médiatiques ? **12.** La radio a-t-elle un/des avantage(s) vis-à-vis d'autres médias (presse, TV…) sur les enceintes connectées ? Quel est/sont-il(s) ? **13.** Sur les enceintes connectées, une voix artificielle encadre l'expression radiophonique, ou bien s'y substitue dans le cas des articles écrits lus à l'oral, prisés des titres des presse sur ce support. Le mode d'expression radiophonique, axé sur la voix humaine, risque-t-il de perdre en authenticité (et, dans un contexte publicitaire, en efficacité) ?

CONCLUSION	**Projection**	**14.** Qu'est ce que ce type d'innovations, qui renouvelle le rapport au média audio, pourrait présager des évolutions médiatiques et de la création publicitaire ?

Annexe 12 — Retranscription de l'entretien de Pierre Laurent, chef de produit marketing radio au groupe RTL, en charge du Laboratoire Sonore.

Les pages 187 à 191 ont été retirées de la version diffusée en ligne.

Annexe 13 — Guide d'entretien de Jérémy Valet, co-directeur de M6 Unlimited.

THÈME 1 PRÉSENTATION	**Profession**	**1.** Pourriez-vous vous présenter (profession, parcours) ? **2.** Votre métier provient-il d'un goût pour la radio, et plus précisément le groupe radio RTL ? **3.** Vous avez récemment changé de fonctions, passant des OPS Radio à M6 Unlimited, structure visant à proposer des solutions crossmédias aux annonceurs. Qu'est ce qui a motivé ce choix, et qu'est ce que ce type de structure indique des évolutions du secteur des médias ?
THÈME 2 LE TRAVAIL DU SON ET DE LA VOIX DANS UN CONTEXTE RADIOPHONIQUE	**Qualité vocale et sonore du média radio**	**4.** Qu'est ce qui fait selon vous la singularité du média radio ? **5.** Comment qualifieriez-vous la relation qui se noue entre des auditeurs et les animateurs à l'antenne ? **6.** Quelles sont les propriétés et les effets de la voix au sein de ce dispositif radiophonique ? **7.** D'après un baromètre réalisé chaque année par Kantar pour mesurer la confiance des Français dans les médias, la radio se place en première position quant à la fiabilité présupposée et ressentie de ses informations (56% contre 52% pour la presse et 48% pour la TV). Est ce pour les mêmes raisons que ce nous venons d'indiquer ?
	Le contexte radiophonique et son évolution	**8.** L'ajout d'image à la radio, notamment sur les réseaux sociaux, altère-t-elle ce qui fait la spécificité de la relation qu'elle instaure avec les publics ? **9.** Les contenus sonores délinéarisés remettent en cause une autre dimension du dispositif radiophonique : une grille structurant un flux en direct. Quelles menaces y a-t-il ici pour la radio, et quelles solutions met-elle en place pour y remédier ?
THÈME 3 LES ENCEINTES CONNECTÉES	**Ensemble des médias présents sur les enceintes connectées**	**10.** Que pensez-vous du mode d'interaction permis par les enceintes connectées ? Y a-t-il un nouveau champ de création médiatique qui se déploie ? **11.** Les fonctions les plus utilisées des enceintes connectées sont les questions d'ordre général sur internet, le minuteur et la musique. Quel est l'intérêt pour un média d'investir dans ce champ ? **12.** Qu'observez-vous à votre niveau quant au positionnement des médias sur les enceintes connectées : leurs discours marketing, la manière dont elles légitiment leur présence ?

	Place de la radio d'un point de vue éditorial	**13.** Sur les enceintes connectées, une voix artificielle encadre ou génère les contenus dans le cas du *text to speech*, prisé par les titres de presse. La radio pourrait-elle apparaître comme plus authentique sur ce support grâce à son travail de voix naturelles ? **14.** Par extension, pensez-vous que la radio a une légitimité particulière sur les enceintes connectées, issue de son identité de média ? **15.** Les larges audiences de la radio, structurellement vieillissantes, lui confèrent un poids historique et économique. Sont-elles cependant une contrainte sur un support où le dynamisme et l'innovation semblent être de rigueur ? **16.** En termes de monétisation, la publicité classique, le sponsoring et les contenus de marque cohabitent sur ce support, très empreint des codes de l'audio numérique. Un développement en OPS est-il envisageable sur les enceintes connectées ?
CONCLUSION	**Projection**	**17.** Que pensez-vous de l'analyse qui est parfois faite d'une probable fonte de la radio dans le champ plus large de l'audio numérique ? **18.** Les développements de la radio sur les enceintes connectées sont-ils un indice d'une quête plus large de renouvellement pour ce média ? **19.** Dans un article publié dans *Les Échos* (« Les médias parient sur les enceintes connectées », 13 juin 2018 – n°2215), Charles-Emmanuel Bon, ancien Directeur du Développement du pôle radio RTL, fait la déclaration suivante : « Les enceintes connectées pourraient devenir le relais de la radio dans le futur ». Que veut-il dire selon vous ? **20.** Avec les enceintes connectées, la machine va jusqu'à remplacer la voix humaine (non sans problèmes techniques pour l'instant). Qu'est ce que cela présage des futures évolutions des contenus et des usages médiatiques ?

Annexe 14 — Retranscription de l'entretien de Jérémy Valet, co-directeur de M6 Unlimited.

Les pages 194 à 204 ont été retirées de la version diffusée en ligne.

Annexe 15 — Guide d'entretien de Guillaume Duhamel, concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL.

THÈME 1 PRÉSENTATION	**Profession**	**1.** Pourriez-vous vous présenter (profession, parcours) ? **2.** Votre métier provient-il d'un goût pour la radio, et plus précisément le groupe radio RTL ?
THÈME 2 LE TRAVAIL DU SON ET DE LA VOIX DANS UN CONTEXTE RADIOPHONIQUE	**Qualité vocale et sonore du média radio**	**3.** Qu'est ce qui fait selon vous la singularité du média radio ? **4.** Comment qualifieriez-vous la relation qui se noue entre des auditeurs et les animateurs à l'antenne ? **5.** Quelles sont les propriétés et les effets de la voix au sein de ce dispositif radiophonique ? **6.** D'après un baromètre réalisé chaque année par Kantar pour mesurer la confiance des Français dans les médias, la radio se place en première position quant à la fiabilité présupposée et ressentie de ses informations (56% contre 52% pour la presse et 48% pour la TV). Est ce pour les mêmes raisons que ce nous venons d'indiquer ?
	Le contexte radiophonique et son évolution	**7.** L'ajout d'image à la radio, notamment sur les réseaux sociaux, altère-t-elle ce qui fait la spécificité de la relation qu'elle instaure avec les publics ? **8.** Les contenus sonores délinéarisés remettent en cause une autre dimension du dispositif radiophonique : une grille structurant un flux en direct. Quelles menaces y a-t-il ici pour la radio, et quelles solutions met-elle en place pour y remédier ?
THÈME 3 LES ENCEINTES CONNECTÉES	**Ensemble des médias présents sur les enceintes connectées**	**9.** Que pensez-vous du mode d'interaction permis par les enceintes connectées ? Y a-t-il un nouveau champ de création médiatique qui se déploie ? **10.** Les fonctions les plus utilisées des enceintes connectées sont les questions d'ordre général sur internet, le minuteur et la musique. Quel est l'intérêt pour un média d'investir dans ce champ ? **11.** Qu'observez-vous à votre niveau quant au positionnement des médias sur les enceintes connectées : leurs discours marketing, la manière dont elles légitiment leur présence ?

	Place de la radio d'un point de vue éditorial	**12.** Sur les enceintes connectées, une voix artificielle encadre ou génère les contenus dans le cas du *text to speech*, prisé par les titres de presse. La radio pourrait-elle apparaître comme plus authentique sur ce support grâce à son travail de voix naturelles ? **13.** Par extension, pensez-vous que la radio a une légitimité particulière sur les enceintes connectées, issue de son identité de média ? **14.** Les larges audiences de la radio, structurellement vieillissantes, lui confèrent un poids historique et économique. Sont-elles cependant une contrainte sur un support où le dynamisme et l'innovation semblent être de rigueur ? **15.** En termes de monétisation, la publicité classique, le sponsoring et les contenus de marque cohabitent sur ce support, très empreint des codes de l'audio numérique. Un développement en OPS est-il envisageable sur les enceintes connectées ?
CONCLUSION	**Projection**	**16.** Que pensez-vous de l'analyse qui est parfois faite d'une probable fonte de la radio dans le champ plus large de l'audio numérique ? **17.** Les développements de la radio sur les enceintes connectées sont-ils un indice d'une quête plus large de renouvellement pour ce média ? **18.** Dans un article publié dans *Les Échos* (« Les médias parient sur les enceintes connectées », 13 juin 2018 – n°2215), Charles-Emmanuel Bon, ancien Directeur du Développement du pôle radio RTL, fait la déclaration suivante : « Les enceintes connectées pourraient devenir le relais de la radio dans le futur ». Que veut-il dire selon vous ? **19.** Avec les enceintes connectées, la machine va jusqu'à remplacer la voix humaine (non sans problèmes techniques pour l'instant). Qu'est ce que cela présage des futures évolutions des contenus et des usages médiatiques ?

Annexe 16 — Retranscription de l'entretien de Guillaume Duhamel, concepteur-rédacteur aux Opérations Publicitaires Spéciales du groupe radio RTL.

Les pages 207 à 213 ont été retirées de la version diffusée en ligne.

Annexe 17 — Textes des contenus de présentation des offres de médias associés au lancement d'Amazon Alexa en France, publiés entre le 1er et le 30 juin 2018.

Les pages 214 à 219 ont été retirées de la version diffusée en ligne.

Annexe 18 — Tableau comparatif des offres des stations de radio françaises sur Amazon Echo au 12 août 2018 (résultats à la requête « radio » sur amazon.fr/skills).

Radio	Skill	Description	Direct	Délinéarisé	Contenus dédiés
Sud Radio	Sud Radio	Cette Skill vous permet d'écouter Sud Radio en direct sur Alexa. Sud Radio, Parlons Vrai, avec Brigitte Lahaie, Patrick Roger, Natacha Polony, Liane Foly, Valérie Expert, André Bercoff, Jacques Pessis, Mickael Darmon, Philippe David, Christophe Bondet, Yolaine de la Bigne, Natacha Harry, Laurence Perault, Julien Martinez, Pauline Dubuisson, Corinne Masiero, Philippe Vandel, Christophe Malavieille, Arthur Jugnot, Julien Martinez, François Trillo, Daniel Herrero... Sud Radio, Parlons Vrai ! La radio généraliste et nationale.	Oui	Non	Non
Fun Radio	Fun Radio	Retrouvez tout Le son Dancehour sur Alexa. Cette skill vous permet d'écouter votre radio en live, mais aussi d'écouter vos programmes favoris à tout moment et à la journée. A tout de suite sur Fun Radio !	Oui	Oui	Non
France Bleu	France Bleu	France Bleu vous permet d'écouter la radio en direct, de suivre en continu les actus de votre région, les infos de proximité, toutes les actualités de votre radio FM locale !	Oui	Non	Non
France Bleu	France Bleu Santé et Bien-Etre	Dans France Bleu - Santé et bien être, retrouvez : "Le conseil santé de France Bleu" par Vanessa Lambert. Un conseil santé quotidien pour prendre soin de soi.	Non	Oui	Non
	France Bleu Sciences	Dans France Bleu - Sciences, retrouvez l'émission "On n'arrête pas le progrès". Un voyage à travers les temps modernes, sur le des des inventions et créations qui ont marqué notre vie quotidienne.	Non	Oui	Non
Virgin Radio	Virgin Radio	Avec la skill Virgin Radio, retrouvez tout l'univers Pop Rock Electro Virgin Radio. Écoutez la radio en direct, Ecoutez vos émissions préférées : « Virgin Tonic » avec Camille Combal, « Le Lab Virgin Radio » avec Lionel Virgille, Cauet S' Lâche, Electroshock....	Oui	Oui	Non
France Musique	France Musique	Écoutez France Musique en direct, la radio de toute la musique classique et du jazz. Retrouvez toute l'actu musicale, les sorties de disques classiques ou jazz, les spectacles, les opéras ou les comédies musicales incontournables. Découvrez les 7 webradios jazz, musiques classiques et musique contemporaine, les concerts des orchestres de Radio France, musiques du monde et musiques de films ou comédies musicales. Zappez facilement d'une musique à l'autre, d'une webradio à l'autre pour découvrir la richesse des sélections et des émissions de France Musique !	Oui	Non	Non
Mouv'	Mouv	Découvrez Mouv', radio hip-hop rap et sa webradio musicale. "Alexa, lance Mouv'" - retrouvez Mouv' en direct : la radio hip-hop, rap, rnb et profitez des meilleurs titres hip-hop du moment, allant du rap à l'électro en passant par le rnb et le reggae. "Alexa, demande à Mouv' d'écouter 100% Mix" - webradio hip-hop et rap, exclusivement numérique, découvrez avec cette radio en ligne la playlist de tous les tracks diffusés sur Mouv, sans pub et sans blabla. Retrouvez gratuitement tous les programmes de Mouv' avec Pascal Cetran, Caroline, First Mike, Mustafa El Atrassi, Willaxxx, Selecta K-za et Lise Pressac, Olivier Cachin, Francis Vied, R-Ash, Supa, Dirty Swift, La Caution, Ilan Malka, Juan Massenya, Akhenaton, Christophe Crenel, Sandra, Greg, T-Miss, Guillaume Tatu, Damien et plus encore... Mouv' et la webradio 100% Mix vous sont proposées en audio de haute qualité : profitez du meilleur de votre enceinte grâce au format AAC. A tout de suite à l'écoute de Mouv' !	Oui	Oui	Non
RTL2	Radio RTL2	Retrouvez tout Le son Pop-Rock sur Alexa. Cette skill vous permet d'écouter votre radio en live, mais aussi d'écouter et réécouter vos programmes favoris à tout moment de la journée. A tout de suite sur RTL-2 !	Oui	Oui	Non
FIP	FIP	Découvrez l'univers musical de Fip, la radio la plus éclectique du monde ! Ecoutez Fip en direct et ses 7 webradios rock, électro, jazz, groove, musiques du monde, reggae et nouveautés. Zappez facilement d'une musique à l'autre, d'une webradio à l'autre pour découvrir la richesse des sélections de Fip radio ! Toutes nos playlists sont confectionnées à la main par l'équipe de programmateurs de Fip, depuis Paris, garanties sans publicité. Vous profiterez d'une grande variété de styles musicaux pour voyager, danser, travailler ou le famiente : jazz, musiques du monde, groove, pop, rock, blues, chanson française, musiques électroniques, classique, électro, bandes originales de films.	Oui	Non	Non
France Culture	France Culture	Écoutez France Culture en direct et retrouvez le regard sur notre société de la radio de tous les savoirs, découvrez ses émissions et ses thématiques phares : culture, philosophie, information, littérature, fictions radiophoniques, documentaires et reportages, idées, arts et spectacles, religions, histoire, sciences. Annoncez "Alexa, lance France Culture" pour retrouver votre radio en direct sur Alexa. Vous pourrez aussi accéder directement à la dernière émission, au dernier épisode d'une sélection de rendez-vous parmi les plus populaires de Culture.	Oui	Oui	Non
	France Culture - Actualités Culturelles	Dans France Culture - Actualités culturelles, retrouvez l'émission "L'édito Culturel".	Non	Oui	Non

Franceinfo	franceinfo	Suivez toute l'actualité en temps réel avec le direct radio de franceinfo, l'offre d'information en continu du service public. Gardez le réflexe info avec la radio franceinfo au rythme des journaux, interviews, débats et reportages. Annoncez "Alexa, lance franceinfo" pour profiter de franceinfo en direct sur Alexa. Vous pourrez aussi accéder directement à la dernière chronique, au dernier épisode d'une sélection de rendez-vous d'info et de chroniques quotidiennes parmi les plus populaires de franceinfo	Oui	Oui	Non
	Le Flash franceinfo	Ce flash est votre condensé d'actus quotidiennes en 2min ! Ecoutez toutes les dernières actus politiques, économiques, internationales et sportives. Mis à jour plusieurs fois par heure, vous ne passerez plus à côté d'une actualité importante !	Non	Oui	Non
	franceinfo - hightech	Dans franceinfo - High Tech, retrouvez l'émission "Le nouveau Monde" : Jérôme Colombain s'intéresse aux innovations, aux produits high-tech, aux nouvelles technologies. Eclairage sur le monde virtuel et technologique.	Non	Oui	Non
	franceinfo - économie	Dans franceinfo - Economie, retrouvez l'émission "Le Décriptage Eco" : Chaque jour, Fanny Guinochet, journaliste de franceinfo et à l'Opinion propose son regard sur un événement marquant de l'actualité économique du jour.	Non	Oui	Non
	franceinfo - sports	Dans franceinfo - Sports, retrouvez l'émission "Le Journal des sports".	Non	Oui	Non
France Inter	France Inter	Écoutez France Inter en direct, retrouvez la matinale la plus suivie de France et toutes vos émissions favorites ! Écoutez les journaux et reportages, les interviews politiques et les chroniques humoristiques. Suivez l'actualité culturelle et musicale, les critiques de spectacles, retrouvez la playlist de France Inter… Découvrez tout l'univers de la radio avec l'appli officielle. Annoncez "Alexa, lance France Inter" pour profiter de France Inter en direct sur Alexa. Vous pourrez aussi accéder directement à la dernière émission, au dernier épisode d'une sélection de rendez-vous parmi les plus populaires d'Inter	Oui	Oui	Non
	France Inter - Humour	Pendant l'été, retrouvez les humoristes de France Inter chaque jour à 8h55	Non	Oui	Non
	France Inter - Géopolitique	Dans France Inter - Géopolitique, retrouvez l'émission "Géopolitique" : Le regard de Bernard Guetta sur la politique internationale.	Non	Oui	Non
	France Inter - Hightech	Dans France Inter - High tech, retrouvez l'émission "C'est déjà demain" : Plongée avec Hélène Chevallier dans les innovations qui vont changer nos vies. Au-delà des technologies, l'avenir c'est aussi mieux consommer, plus partager, moins polluer.	Non	Oui	Non
Chérie	Chérie	Avec la Skill Chérie, écoutez en direct et en illimité plus de 40 min de Pop Love music Non Stop ! A tout moment, accédez à plus de 35 radios digitales selon votre envie et retrouvez tous vos artistes Pop Love Préférés. Pour lancer une radio digitale précise dites : "Alexa, lance chérie." par exemple. Ou laissez Alexa choisir pour vous : "Alexa, demande à chérie de lancer une radio au hasard." Vous pouvez choisir d'écouter un genre musical parmi les choix suivants : Love, Pop, Relax, Ambiance, Sport, Décennies. Pour cela dites : "Alexa, demande à chérie de mettre un genre musical." Votre musique préférée, c'est sur Chérie Pop Love Music !	Oui	Non	Non
Nova	La Revue de Presse Musicale de Nova	Chaque jour, Nova se pose en vigie des dernières infos musicales. En quelques minutes, vous savez tout sur les albums en préparation, les sorties attendues, les festivals partout en France… On vous propose aussi régulièrement des inédits, des écoutes en avant-première et des surprises.	Non	Oui	Non
NRJ	NRJ	Avec la Skill NRJ, écoutez toutes vos émissions en DIRECT, écoutez les plus gros Hits du moment et découvrez les meilleures nouveautés ! A tout moment, accédez à plus de 50 radios digitales Dance, RnB, Rap, Electro ou Pop … et retrouvez tous vos artistes préférés. Pour lancer une radio digitale précise dites : "Alexa, lance NRJ HITS" par exemple ou laissez Alexa choisir pour vous : "Alexa, demande à NRJ de lancer une radio au hasard" Vous pouvez choisir d'écouter un genre musical parmi les choix suivants : HITS, DANCE, ELECTRO, RAP, RnB, AMBIANCE, SPORT. Pour cela dites : "Alexa, demande à NRJ de mettre un genre musical" Votre radio préférée, vos hits préférés : ici c'est NRJ !	Oui	Oui	Non
RFM	RFM	Avec la Skill RFM écoutez en direct le meilleur de la musique des années 80 à aujourd'hui. RFM c'est également tous les matins de la bonne humeur, des bons plans et des cadeaux avec « Le meilleur des réveils » présenté par Elodie Gossuin et Albert Spano.	Oui	Non	Non
Europe 1	Europe 1	Avec la skill Europe 1, vous pouvez écouter tous les programmes en direct, retrouvez une sélection d'émissions en podcast et demander le dernier journal d'information de la rédaction. Retrouvez la revue de presque de Nicolas Canteloup, Au coeur de l'histoire de Franck Ferrand, les récits de Christophe Hondelatte, la matinale de Patrick Cohen, la chronique foot de Julien Cazarre, le coup de patte de Nadia Daam ou encore l'actualité numérique avec Grégoire Martinez. N'hésitez pas à demander à Alexa la liste de tous les podcasts disponibles.	Oui	Oui	Non
Skyrock	Skyrock	Avec la skill Skyrock, écoute les meilleurs sons du moment et retrouve toutes tes émissions préférées ! Ecoute le Live, ou ambiance-toi sur les meilleurs sons en découvrant nos 8 autres radios Skyrock. A tout moment pendant ton écoute, utilise la commande « Alexa, suivant » ou « Alexa, précédent » pour naviguer entre les différentes radios.	Oui	Oui	Non

RTL	RTL	Editions d'information, Grosses Têtes… Retrouvez tous les programmes de la première radio de France sur Alexa. Cette skill vous permet d'écouter RTL en live et aussi d'écouter et réécouter toutes vos émissions préférées, à tout moment de la journée. À tout de suite sur RTL !	Oui	Oui	Non
	Journal RTL	"Les équipes de la rédaction RTL travaillent 7 jours sur 7 et 24 heures sur 24 pour vous informer en temps réel et décrypter l'actualité pour mieux la comprendre. Retrouvez toutes les heures, la dernière édition du journal audio RTL, ainsi que la météo à 7 jours."	Non	Oui	Non
RFI	RFI Monde	Retrouvez l'ensemble des dernières informations politiques françaises et internationales, culturelles, sportives dans le journal de RFI.	Non	Oui	Non
	RFI	Avec l'application RFI spécialement conçue pour Alexa, il n'a jamais été aussi simple d'écouter et d'interagir avec vos contenus RFI préférés. Vous pouvez écouter le direct de RFI "Monde" ou "Afrique", ainsi que les derniers journaux associés dont le "journal en français facile". Une sélection avec vos émissions préférées de RFI est à portée de « voix ». De plus, si vous souhaitez écouter les infos autour d'un pays en particulier, dites « Alexa, ouvre RFI et cherche les dernières infos du Mali » et vous pouvez obtenir les dernières nouvelles du pays qui vous intéresse.	Oui	Oui	Non
	RFI - Journal en français facile	Le journal en français facile se caractérise par des phrases simples afin que les auditeurs souhaitant apprendre le français ou maîtrisant mal cette langue puissent le comprendre aisément. Qu'est ce qu'un « droit de veto » ? À quoi sert un « négociateur » ? « L'apartheid », en Afrique du Sud, c'était quoi ? Pourquoi parle-t-on de « l'élargissement » de l'Union européenne ? Ces mots, vous les entendez souvent dans les journaux. Pour comprendre les mots et les concepts, un vrai journal d'information qui raconte chaque événement avec son contexte, son histoire et ses personnages.	Non	Oui	Non
RMC	L'actualité Sport avec RMC	Le Flash Sport de RMC SPORT vous propose plusieurs fois par jour les dernières actualités Football (Ligue 1, Premier League, Champions League...), Rugby, Tennis et tous autres les sports. Après avoir installé l'application, dites « Alexa, quel est mon flash quotidien » ou « Alexa, quelles sont les infos ? » pour écouter le flash.	Non	Oui	Non
Rire et Chansons	La Blague du Jour avec Rire et Chansons	Comment bien commencer votre journée ? En écoutant la blague du jour de Rire & Chansons ! Tous les matins, la skill Rire & Chansons vous propose la dose de bonne humeur à télécharger sur «Amazon Market» !	Non	Oui	Non
Les Indés Radios	Les Indés Radios	Radios locales et radios thématiques, avec Les Indés Radios la radio indépendante vit plus fort. Le Groupement Les Indés Radios réunit 130 radios FM, écoutées chaque jour par plus de 8,6 millions d'auditeurs ! Vous pouvez demander de lancer plus de 130 stations comme Latina, Alouette, Chante France, Radio FG, Hit West, Contact FM.	Oui	Non	Non
			TOTAL	**TOTAL**	**TOTAL**
			35 éléments	35 éléments	35 éléments
			19 oui	28 oui	0 oui
			16 non	7 non	35 non

Résumé.

Notre recherche entend déterminer la place de la création radiophonique, incarnée par des stations diffusant historiquement sur la bande FM, au sein d'une économie numérique dont les enceintes connectées cristallisent les tensions. Ces transistors augmentés de la possibilité d'une interaction, proposant des fonctionnalités multiples allant de la domotique à la consommation de contenus médiatiques, déplacent le travail de la voix d'un monopole des stations de radio vers une concurrence élargie à tous types d'entreprises médiatiques et des Intelligences Artificielles qui altèrent l'origine corporelle de ce matériau. Par le biais d'une analyse alternant des aspects épistémologiques, économiques et technologiques, ce mémoire dessine le paysage numérique dans lequel les radios évoluent, supposant de remédier leurs identités et leurs pratiques.

La voix apparaît comme la conjugaison de l'expression sensible d'un corps projeté dans le monde et d'un arbitraire du signe linguistique qui, tout en rendant le contenu d'un message intelligible à tous, en amenuise la portée émotionnelle. Le maniement de cette tension entre la musicalité d'un grain de la voix individuel et l'aspect conventionnel d'une grammaire qui fait société est au fondement de la création radiophonique, qui en transportant la voix dans des ondes met en œuvre des techniques destinées à en préserver l'authenticité. Cette dernière, caractéristique de l'écoute radiophonique, se décompose en des effets d'intimité, de communauté, de réel et d'imaginaire d'autant plus prégnants qu'ils respectent le dispositif initial que construit la radio, acousmatique et en direct. La virtualité numérique, qui consiste en la construction d'un espace non plus caractérisé par le transport mais par le déplacement de la source de la voix, en réplique les propriétés d'interaction (des dialogues s'instaurent entre des utilisateurs et des machines) et d'immersion (elle constitue une seconde réalité, aussi enveloppante que l'écoute d'un son à la source invisible). Les enceintes connectées, qui mobilisent la voix dans la virtualité numérique, qu'elles intègrent dans la réalité physique — le foyer de leurs utilisateurs — constituent autant un aboutissement de la virtualité comme prolongement du réel physique qu'une menace pour les radios, dont elles déplacent la source corporelle vers des voix artificielles.

Les enceintes connectées mettent à jour un champ concurrentiel d'entreprises médiatiques dominé par les GAFAM (Google, Amazon, Facebook, Apple et Microsoft), dont le degré d'implantation sur un territoire conditionne la richesse de l'offre de contenus proposés aux utilisateurs. La France est en cela un marché embryonnaire, dont les acteurs médiatiques ont les yeux rivés vers les États-Unis, pays où cette technologie imprègne de plus en plus les usages. Les

concepteurs des enceintes connectées, afin de conquérir les foyers du monde entier, emploient des recherches technologiques importantes pour intégrer les assistants vocaux aux modes de vie des usagers. Cela n'apparaît possible qu'en conjuguant intelligence et émotion artificielles, la seconde n'étant aujourd'hui qu'un fantasme et la première une réalité nuancée par la forte incorporation de travail humain dans ses processus de conception, brisant un idéal d'automaticité.

Dans ce contexte, les stations de radio sont dans une position ambivalente. Elles bénéficient d'une avance et d'une légitimité car le travail de la voix est ancré dans leurs identités, mais pâtissent d'une économie numérique qui contraint leur forme (avec l'ajout d'image et un souci de viralité) et abaisse leurs revenus (de par la multiplication de canaux d'expression et la cannibalisation de leurs apports monétaires par les possesseurs des plateformes). Ce dernier point est le plus sensible, car notre recherche semble souligner que la radio ne pourra concrétiser son avantage concurrentiel qu'en incorporant des compétences technologiques aux savoir-faire radiophoniques, supposant des ressources à la fois financières et humaines. La radio, aux revenus plus modestes que d'autres médias, et dont les professionnels ne lient pas tous ces enjeux aux pratiques de leurs métiers, pourrait par conséquent voir ses stations historiques fondre dans le champ de l'audio numérique. Cette supposition ne trouvera son élucidation que dans la mise en place des stratégies d'une multiplicité d'acteurs qui demeure en cours, et dont notre recherche propose une photographie à un moment clé.

En adoptant le prisme de la voix pour analyser la place de la radio sur les enceintes connectées, notre recherche suit le déplacement du corps par la technique — de l'air expiré aux ondes radio, du micro aux voix des Intelligences Artificielles —, qui redéfinit les modes de réception des publics et les pratiques professionnelles de l'ensemble des acteurs de la création médiatique.

Mots-clés.

Voix — Radio — Enceintes connectées — Médias — Numérique — Intelligence Artificielle