

## **Traduire à voix haute:** la traduction dictée interactive comme solution ergonomique

By Julián Zapata & Elizabeth C. Saint (University of Ottawa, Canada)

### **Abstract & Keywords**

#### **English:**

Voice recognition (VR) technology has greatly evolved since computer tools first came into our lives. It has been increasingly adopted by translators, who turn to the technology to minimise or prevent certain physical or mental health problems, among other reasons. VR has also helped bring translation dictation back into the profession. As shown in the first study presented in this article, the ergonomic advantages of using VR are genuinely valued by translators. However, we argue that VR could be improved by combining it with the various input modes offered by newer multimodal and mobile interfaces, which offer many advantages. We therefore advocate moving beyond translation dictation to develop an interactive translation dictation (ITD) environment, through which the professional translator's workstation would become a natural extension of his or her capabilities. Our second study presents the results of an experiment designed to assess the ergonomic gains of using a prototypical ITD environment.

#### **French:**

Les technologies de reconnaissance vocale (RV) se sont largement développées depuis l'arrivée des outils informatiques dans notre vie. De plus en plus employée par les traducteurs pour, entre autres, réduire ou prévenir certains problèmes de santé physiques ou psychologiques, la RV a aussi contribué au retour de la traduction dictée chez ces professionnels. Comme la première étude présentée dans cet article le montre, les avantages ergonomiques de l'utilisation de la RV sont véritablement appréciés par les traducteurs. Nous avançons toutefois que la RV gagnerait à être combinée aux multiples atouts que présentent les interfaces multimodales des outils technologiques émergents et mobiles. Plus que de traduction dictée, il s'agirait alors d'offrir un espace de traduction dictée interactive (TDI), par lequel la plateforme de travail du traducteur professionnel deviendrait un prolongement de ses compétences. Notre seconde étude présente les résultats d'une expérience d'utilisation d'un prototype d'environnement de TDI visant à en évaluer sa valeur ajoutée en matière d'ergonomie.

**Keywords:** voice recognition, multimodal interaction, interactive translation dictation, translation technology, ergonomics, reconnaissance vocale, interaction multimodale, traduction dictée interactive, traductique, ergonomie

---

### **Introduction**

Depuis les années 1950 et l'arrivée de l'ordinateur, les chercheurs[1] de l'informatique travaillent à développer des moyens de traiter automatiquement le langage humain. Leurs travaux ont donné lieu à la conception de divers systèmes employant les technologies de reconnaissance et de synthèse vocales, tels que les agents conversationnels Alexa, Cortana ou Siri. Ces deux technologies permettent à la machine de « reconnaître » automatiquement ce que dit une personne (reconnaissance vocale) et de générer automatiquement de la parole en simulant la voix humaine (synthèse vocale). La première s'est améliorée plus lentement que la seconde, mais on constate que les technologies de reconnaissance vocale (RV) arrivent enfin à maturité et qu'elles sont maintenant employées avec succès dans des domaines aussi divers que les services à la clientèle, le soutien technique téléphonique et en ligne, les systèmes de navigation des véhicules ou encore les systèmes d'exploitation des ordinateurs personnels, tablettes ou téléphones intelligents (Jurafsky et Martin 2009 : 285). Ces avancées coïncident avec le développement des systèmes informatiques multimodaux qui rendent l'interaction avec la machine plus naturelle et plus ergonomique[2].

Il n'est pas étonnant que la traduction profite, elle aussi, de ces progrès technologiques; d'autant que l'idée de jumeler la RV et la traduction dictée (Brousseau et coll. 1995; Brown et coll. 1994; Dymetman et coll. 1994) remonte aux années 1990. Toutefois, les imperfections de la RV à cette époque et son incapacité à automatiser des tâches langagières essentielles, comme la transcription de la dictée, ont contribué à ce que la recherche se tourne plutôt vers la conception d'outils de traductique, capables de prendre en charge des tâches linguistiques périphériques et d'améliorer la performance et l'efficacité des traducteurs (Bowker 2002). Ainsi, ce n'est réellement qu'au début des années 2000 que la RV a commencé à faire partie de l'outillage des traducteurs et, rapidement, elle a présenté de nombreux avantages, notamment ergonomiques. La valeur ajoutée de la RV dans la pratique professionnelle du traducteur justifie que l'on vise son intégration au sein d'une interaction multimodale avec la machine. C'est du moins la proposition que fait cet article en présentant la traduction dictée interactive (TDI), concept innovant introduit par Zapata (2012; 2016) pour l'amélioration des conditions de travail des traducteurs professionnels et de leur productivité.

Nous commencerons par faire un rapide retour historique sur l'intégration de la RV dans la pratique de la traduction et nous présenterons brièvement les résultats d'une étude qui révèlent les avantages que des traducteurs disent tirer de la RV lorsqu'ils l'utilisent pour traduire. Ceci nous amènera à exposer les caractéristiques de l'interaction multimodale et du concept de TDI. Notre article se conclura sur quelques-uns des résultats d'une expérience menée auprès de 14 traducteurs professionnels qui ont traduit à partir d'un poste de travail traditionnel (avec clavier et souris) et avec un environnement de TDI. Nous limiterons notre discussion aux données qualitatives issues des entretiens semi-dirigés avec les participants, durant lesquels ils se sont exprimés sur l'ergonomie des deux environnements de travail.

### **La recherche d'une traductique ergonomique**

La reconnaissance vocale (RV)[3] se définit comme une « [t]echnologie ayant pour but de permettre à un ordinateur de reconnaître les signaux émis par la voix humaine en vue de la transformer en données numériques » (Office québécois de la langue française 2007). Elle se situe dans le domaine du traitement automatique des langues naturelles (TALN), c'est-à-dire tout ce qui unit l'informatique au langage humain. Dès les prémices du TALN, dans les années 1950, l'un des objectifs était de faire de l'ordinateur un agent conversationnel, un interlocuteur de l'être humain. Toutefois, au début des années 1990, on était encore loin d'arriver à cet idéal et la RV éprouvait encore de nombreuses difficultés à analyser la parole humaine dans toutes ses variétés (accent, intonation, sonorité, vitesse, timbre de la voix, etc.). Dans le courant de cette décennie, de nombreux laboratoires d'entreprises commerciales ont largement investi pour développer des systèmes de RV fonctionnels, pouvant diriger des machines pour des tâches bien spécifiques. Ces avancées, satisfaisantes pour les uns (Rabiner 1997), mais frustrantes pour d'autres (Lippmann 1997), ont suscité un intérêt soudain pour le TALN dans le domaine de la traduction (Brousseau et coll. 1995; Brown et coll. 1994; Dymetman et coll. 1994). Toutefois, la RV d'alors n'étant pas encore assez performante pour répondre aux besoins de cette profession, c'est le traitement de tâches linguistiques périphériques qui a été privilégié et qui a donné naissance aux outils de traductique (Bowker 2002). Il a fallu attendre le début des années 2000 et l'arrivée de logiciels de RV commerciaux plus performants pour que cette technologie refasse son apparition dans la boîte à outils de quelques traducteurs (voir, par exemple, Benis 2002; Seaman 2002; Stroman 2002).

Les chercheurs recommencent alors à s'intéresser à l'intégration de la RV dans la pratique professionnelle de la traduction, notamment pour tester si elle aide le traducteur à gagner en productivité (Désilets et coll. 2008; Dragsted, Hansen et Selsøe Sørensen 2009; Reddy et Rose 2010; Vidal et coll. 2006). Le succès des systèmes de RV dans d'autres domaines, la performance de traitement accrue des ordinateurs intégrant la RV et le besoin urgent de concevoir des outils de traductique ergonomiques, c'est-à-dire qui tiennent compte du facteur humain (O'Brien 2012), ont éveillé un réel intérêt chez les chercheurs en traductologie pour les technologies de RV. Ainsi, certains travaux se sont intéressés à l'usage de la RV dans la formation des traducteurs (Mees et coll. 2013) ou des post-éditeurs (Mesa-Lao 2014), ainsi que dans la pratique du sous-titrage (Romero-Fresco 2011). D'autres ont testé la performance de la RV pour la dictée sur différentes interfaces informatiques (Zapata et Kirkedal 2015) ou pour la post-édition de traductions automatiques (Garcia-Martinez et coll. 2014; Torres-Hostench et coll. 2017; Zapata, Castilho et Moorkens 2017). Finalement, quelques-unes se sont penchées sur la perception qu'ont les traducteurs professionnels des avantages de la RV, notamment, pour leur productivité et pour la qualité de leurs traductions (Ciobanu 2014; 2016).

Ces recherches confirment qu'à l'instar du reste des utilisateurs des technologies, la dépendance du traducteur aux systèmes informatiques est en constante progression. Face au besoin palpable d'interfaces spécifiquement ajustées au processus traductif et aux besoins ergonomiques des traducteurs (Carl, Dragsted et Jakobsen 2011; Taravella et Villeneuve 2011; 2013), un appel à de plus amples études expérimentales a été lancé :

More experimental studies of translator-tool interaction could be carried out using formal usability research methods such as screen recording, eye tracking, and observation, the results of which could then be used by translation technology developers to improve the specifications of tools for the benefit of translators and, ultimately, the end users of those translations. (O'Brien 2012 : 116-7)

Étant donné qu'une technologie trop rigide entrave le travail qu'elle est censée appuyer (Karamanis, Luz et Doherty 2011 : 49), les traductologues sont de plus en plus conscients, qu'en cette ère de traductique, il leur faudra accorder une plus grande place à l'étude du facteur humain (O'Brien 2012). Cela permettra alors de concevoir des interfaces logicielles, plus conviviales et mieux adaptées, à la fois aux besoins ergonomiques et documentaires des traducteurs et à la réalité d'un marché mondial de la traduction qui s'élargit et évolue rapidement (Barabé 2013; Taravella et Villeneuve 2013). À ce propos, la première étude que nous exposons ici vise à extraire les avantages ergonomiques et documentaires que ressentent les traducteurs qui emploient la RV dans leur pratique.

### **Étude 1 : Les avantages de l'utilisation de la RV selon les traducteurs (ergonomie, productivité, qualité)**

Le corpus que nous avons analysé est extrait de trois portails en ligne destinés aux traducteurs et se compose de commentaires publiés par ces derniers, entre janvier 2003 et mars 2012 (fin de la collecte de données), au sujet de la RV. Malgré sa petite taille (9 500 mots), notre corpus a l'avantage de présenter des propos spontanés de traducteurs intéressés par la traduction avec RV. Nous présentons les résultats de façon très globale, mais invitons le lecteur à se référer au travail de Zapata (2012) pour accéder à de plus amples détails.

Les raisons ergonomiques sont les premières évoquées par les traducteurs pour justifier leur utilisation de la RV dans leur pratique professionnelle et elles précèdent celles liées à leurs besoins de gagner en productivité ou en qualité. Ainsi, plusieurs indiquent adopter la RV pour deux raisons : de façon préventive, pour diminuer le stress ou éviter des maladies chroniques, comme des maladies musculo-squelettiques (ex. : le syndrome du canal carpien), des maux de dos, aux épaules, aux poignets ou des problèmes oculaires; ou comme mesure réparatrice après avoir contracté ces troubles. En règle générale, les traducteurs se servent de logiciels de RV pour dicter leur traduction et certains y voient, là aussi, des avantages pour leur santé et leur confort personnel. En effet, la dictée leur permet de s'éloigner de leur écran d'ordinateur et de rester productifs sans avoir à rester assis pendant plusieurs heures.

Outre les bienfaits de la mobilité pour la santé, quelques traducteurs indiquent que le fait de pouvoir utiliser un casque d'écoute sans fil ou de s'éloigner de leur poste de travail leur permet de travailler de façon plus décontractée, tout en garantissant la production d'une traduction de qualité. En effet, après avoir constaté que leur volume de travail dépassait leur capacité de saisie à l'ordinateur, certains traducteurs disent s'être tournés vers la RV et avoir atteint une productivité jusqu'à six fois plus élevée en dictant, plutôt qu'en employant un clavier. Un facteur contributeur évoqué est que l'emploi de la RV faciliterait la concentration du traducteur sur la tâche de traduction uniquement, ce qui, en retour, lui permettrait de soigner la qualité du produit final et le rendrait aussi plus créatif. Une poignée de traducteurs mentionnent regretter que l'utilisation du clavier ait un impact si nuisible sur leur créativité et leur compétence de traduction à proprement parler. En somme, selon eux, l'évaluation de la compétence d'un traducteur ne devrait pas avoir à dépendre autant de sa compétence à saisir (rapidement) sa traduction au clavier d'ordinateur.

Alors que les traducteurs emploient les systèmes de RV pour dicter leurs traductions, peu d'entre eux les utilisent pour envoyer des commandes à l'ordinateur ou pour fureter dans des dossiers ou dans le Web. Pourtant, ceux qui y recourent les trouvent particulièrement utiles pour corriger des erreurs de reconnaissance, surtout qu'en le faisant à voix haute, le logiciel de RV apprend des erreurs et devient du même coup plus performant. Quelques traducteurs démontrent un intérêt à envoyer des commandes à divers logiciels de traductique installés sur leur ordinateur, notamment les gestionnaires de mémoire de traduction, mais reconnaissent qu'ils ne sont pas tous compatibles avec la RV. Finalement, des traducteurs développent une interaction que l'on peut considérer de multimodale (Oviatt 2012) en

alternant entre la RV, le clavier et la souris pour dicter leurs traductions. Cette forme d'interaction leur permettrait de contrer les problèmes de reconnaissance ou d'incompatibilité des logiciels de traductique, par exemple. Ces traducteurs ont donc su trouver un équilibre dans cette alternance et tirer profit de la RV, là où elle fonctionne sans erreur ou presque.

L'utilisation de la RV apporterait donc une nouvelle dimension au flux de travail habituel pour les traducteurs. En assurant son adaptation optimale aux outils de traductique, la traduction dictée (TD) pourrait devenir une TD interactive (TDI) qui amènerait le traducteur à interagir avec la machine de façon multimodale, ergonomique, intuitive ou, tout simplement, « humaine » (*ibid.* : 428).

### **L'interaction multimodale, modèle ergonomique**

Les systèmes multimodaux visent à reconnaître les comportements humains et l'utilisation naturelle du langage, afin de parvenir à interagir le plus intuitivement possible avec la machine. Oviatt (*ibid.* : 405) les définit ainsi :

Multimodal systems process two or more combined user input modes – such as speech, pen, touch, manual gestures, gaze, and head and body movements – in a coordinated manner with multimedia system output. (...) This new class of interfaces aims to recognize naturally occurring forms of human language and behavior, which incorporate at least one recognition-based technology (e.g. speech, pen, vision).

En cela, les systèmes multimodaux remettent en question la pertinence des interfaces graphiques traditionnelles telles que les ordinateurs de bureau, qui font essentiellement usage du clavier et de la souris, en explorant d'autres modes de saisie et de sortie (Girol et coll. 2012; Williamson, Crossan et Brewster 2011).

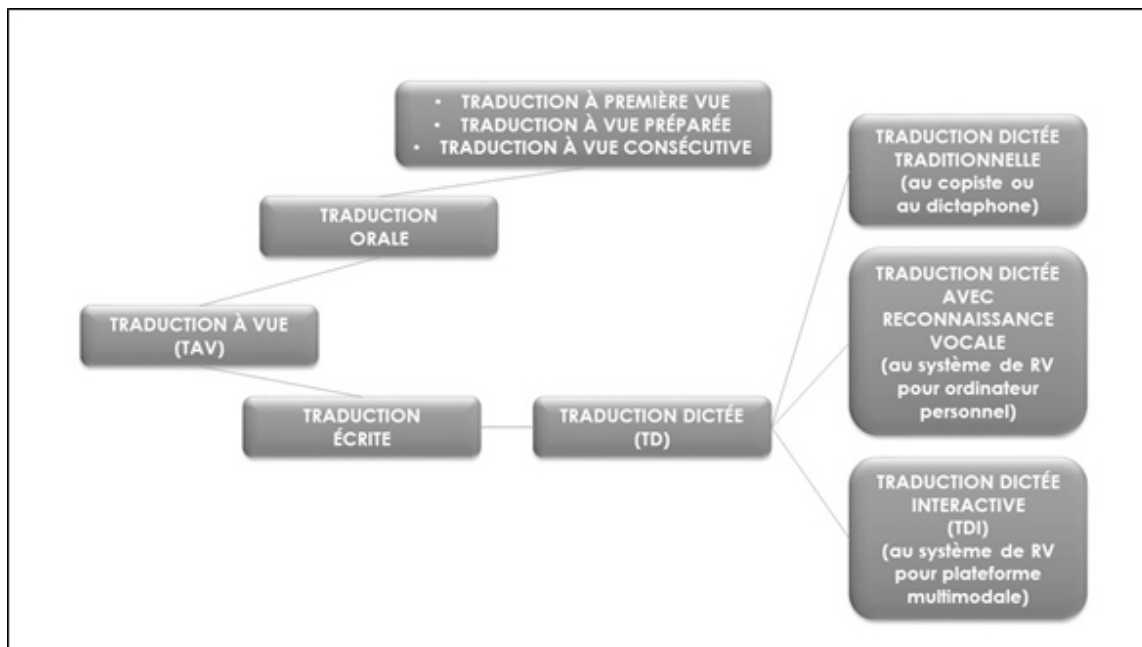
Les avantages de l'interaction multimodale avec la machine sont nombreux, allant d'une meilleure performance (Cohen et coll. 1998) à une flexibilité d'utilisation accrue (usage en extérieur, en déplacement, etc.) (Oviatt et Cohen 2000). C'est aussi ce que révèle une étude de cas menée par Schalkwick et coll. (2010), de l'entreprise Google, sur l'un des exemples actuels les plus concrets et universels d'interface multimodale : le téléphone intelligent. En testant l'usage des commandes vocales pour faire des recherches sur le Web et dans le système d'exploitation du téléphone, les chercheurs ont souligné les nombreux avantages liés à ce type d'interaction, tout en reconnaissant les nombreux défis liés à sa conception. Ainsi, en autorisant la saisie par le toucher et la RV, le téléphone intelligent confirme les gains en mobilité, flexibilité et performance qu'apporte la multimodalité et fournit un modèle d'interaction plus naturelle entre l'humain et la machine. En outre, la polyvalence du système permet à l'utilisateur d'employer simultanément des modes de saisie passifs (clavier, souris) et actifs (voix, doigt, stylet) (Shneiderman 1998 : 318) ou d'alterner entre eux plus aisément, ce qui semble réduire le niveau de frustration des utilisateurs envers la machine, éviter les erreurs (Oviatt et van Gent 1996) et alléger la surcharge cognitive face à des tâches complexes (Oviatt, Coulston et Lunsford 2004).

Même du point de vue de la conception des interfaces, les systèmes multimodaux se sont montrés avantageux, puisque capables de servir des utilisateurs aux profils divers, sans avoir à faire de nombreux compromis pour atteindre l'« utilisateur moyen » (si tant est que celui-ci existe). En somme, les interfaces multimodales ont non seulement le potentiel de fonctionner plus efficacement que les interfaces monomodales, mais elles le font aussi d'une façon plus stable dans des contextes divers et pour des utilisateurs aux profils variés. Les différences comportementales ont d'ailleurs inspiré et guidé un bon nombre des recherches sur la multimodalité. Ainsi, une série d'études empiriques menées par Oviatt, Cohen et Wang (1994) mettent l'accent sur les défis des interfaces multimodales à traiter la variation linguistique des utilisateurs, à l'oral comme à l'écrit. Une autre recherche s'est intéressée aux raisons qui mènent les utilisateurs à interagir différemment avec l'interface multimodale et a tenté de vérifier si elles sont liées à d'autres caractéristiques comportementales (Oviatt, Lunsford et Coulston 2005). En traduction, où les questions de l'ergonomie gagnent en importance, la recherche sur la multimodalité se penche, par exemple, sur les usages de la voix, du stylet ou de l'écran tactile. Pour notre part, nous explorons une approche multimodale de la TD.

### **Au-delà de la dictée : la TDI**

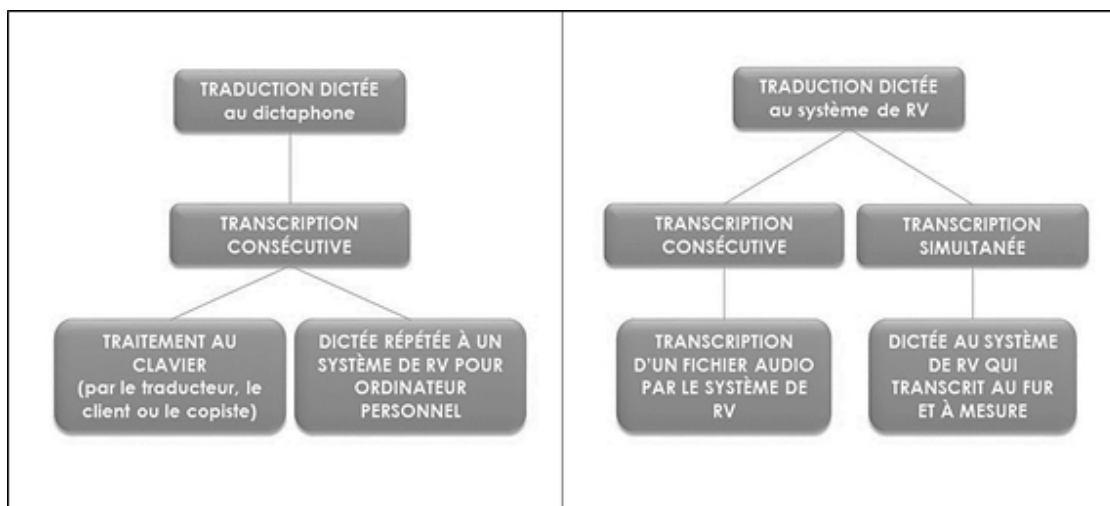
La TDI se définit comme une « technique traductive qui implique l'interaction avec des interfaces multimodales tactiles et vocales, tout au long du processus de traduction, soit la préparation, la production et la révision » (Zapata 2016 : 86, notre traduction). Elle représente une nette évolution des techniques de traduction courantes, comme la traduction saisie à l'ordinateur et la TD, que cette dernière soit traditionnelle avec dictaphone ou qu'elle emploie la RV.

Comme le montre la Figure 1, la TDI constitue une forme de traduction à vue (TAV), technique de traduction à voix haute, employée depuis des décennies dans la formation et la pratique de l'interprétation, mais aussi dans celles de la traduction (Jiménez Ivars et Hurtado Albir 2003).



**Figure 1. Les divers types de TAV**

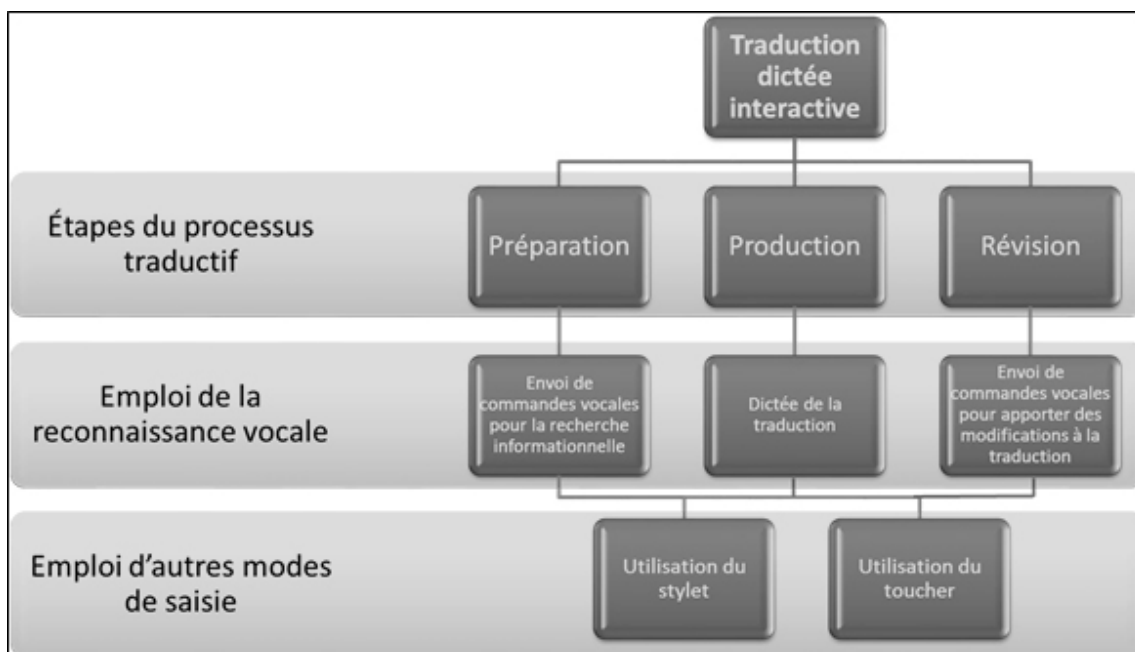
Avant de présenter la TDI plus en détail, et pour mieux comprendre en quoi elle diffère des méthodes de traduction courantes, nous prendrons quelques instants pour décrire ces dernières (traduction saisie à l'ordinateur et autres traductions dictées). Nous excluons de la discussion la traduction dictée directement à un copiste, cette pratique étant, semble-t-il et selon les dires de traducteurs professionnels avec qui nous avons pu nous entretenir, maintenant rare et déconseillée pour des questions de rendement et de rentabilité. Ce qui différencie ces approches réside dans la manière dont la traduction sera transcrite (Figure 2).



**Figure 2. Les modes de transcription des TD, autres que la TDI**

Ainsi, la traduction saisie à l'ordinateur, c'est-à-dire manuellement par l'usage d'un clavier, s'avère moins efficace que les TD pour la concentration du traducteur sur le texte à traduire. En effet, de récentes études (Dragsted et Hansen 2009; Dragsted, Hansen et Selsøe Sørensen 2009; Mees et coll. 2013) ont montré que le fait de produire sa traduction oralement, à la vitesse de la parole, aide le traducteur à se concentrer sur le texte de départ et l'autorise à laisser les questions de transcription et de révision pour plus tard. Cela est particulièrement le cas de la TD au dictaphone et de la TD au système de RV qui seront transcrites de manière consécutive (voir Figure 2). Toutefois, si la TD au dictaphone profite à la concentration sur le texte source, elle présente l'inconvénient d'ajouter la tâche de transcription au traducteur ou l'embauche de personnes complémentaires, à moins, bien sûr, que le client n'accepte de se charger lui-même de la transcription. En ce qui concerne la TD au système de RV avec transcription consécutive, Dragsted, Hansen et Selsøe Sørensen (2009 : 313) ont indiqué qu'elle présentait de plus grands défis d'attention pour le traducteur. De fait, voyant sa traduction s'afficher à l'écran au fur et à mesure qu'il la dicte, il aurait tendance à ne pas résister à la tentation de corriger les erreurs de transcription qui peuvent apparaître. La TD au système de RV avec transcription consécutive vient, sans aucun doute, régler ce problème puisque le traducteur soumet l'ensemble de sa dictée enregistrée à un logiciel de RV qui, à son tour, prend en charge la transcription (Wu 2008). Par contre, cette approche requiert du traducteur qu'il soit équipé d'une variété d'appareils et de logiciels, tels qu'un ordinateur personnel puissant, un dictaphone numérique, un système de RV (unilingue et monomodal) et un logiciel de traitement de textes. Enfin, une autre approche a été envisagée par certains organismes[4] : le travail en équipe de traducteurs dictant des traductions à l'aide de dictaphones numériques et de copistes transcrivant les enregistrements à l'aide de logiciels de RV. D'ailleurs, certains marchands offrent des solutions logicielles permettant ce type de travail collaboratif.

En somme, les TD auront présenté de nombreux avantages sur la traduction saisie à l'ordinateur, mais elles comportent encore de nombreux inconvénients que la TDI pourrait, croyons-nous, estomper en offrant une interaction multimodale et ergonomique entre les traducteurs et la machine (voir Figure 3).

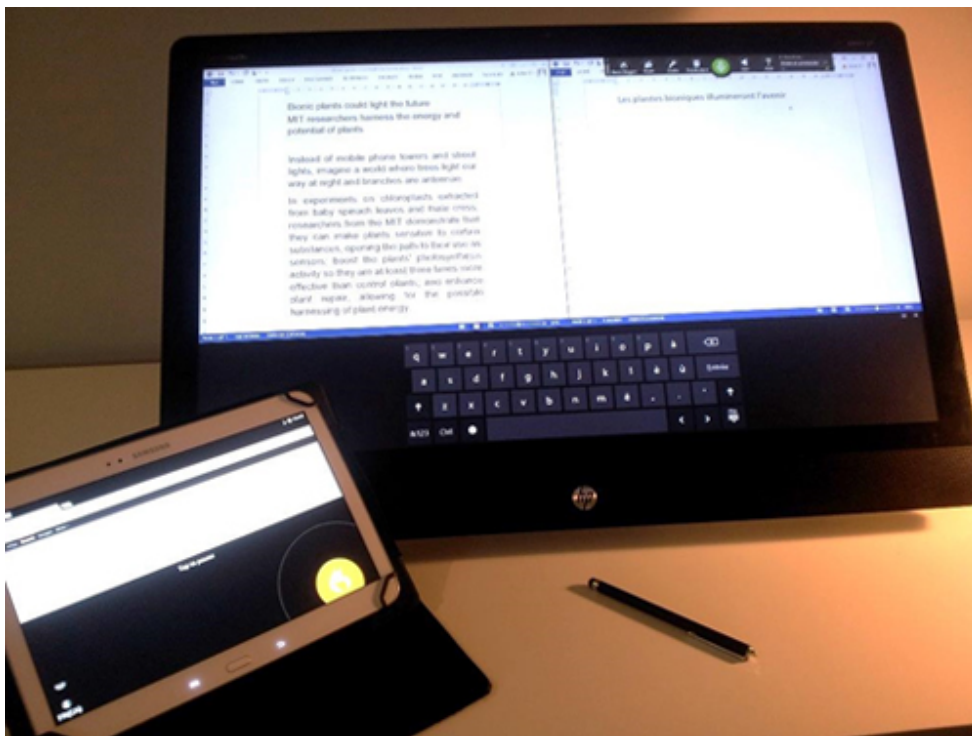


**Figure 3. Description de la TDI dans le processus traductif**

En effet, en permettant au traducteur d'utiliser le système de RV pour faire plus que de la TD (par exemple, de la recherche documentaire et terminologique dans les corpus ou bases de données terminologiques) et d'y ajouter l'usage du stylet ou du toucher quand le besoin se fait sentir, la TDI a le potentiel de devenir « un prolongement naturel du processus de travail habituel du traducteur professionnel » (Zapata et Quirion 2016 : 539). Cela vaut d'autant plus que les interfaces multimodales commerciales à notre disposition (tablettes, ordinateurs à écran tactile) n'ont jamais été aussi robustes et qu'elles ne vont qu'en s'améliorant. La deuxième étude que nous exposons ici vise précisément à tester cette hypothèse.

### Étude 2 : Les avantages de l'utilisation d'un environnement de TDI selon les traducteurs

Lors de cette expérience, 14 traducteurs professionnels qui travaillent de l'anglais vers le français ont eu à effectuer une traduction en interaction soit avec un poste de travail traditionnel (avec clavier et souris) soit avec un environnement de TDI (Figure 4). Ce dernier consistait en deux interfaces multimodales commerciales, à savoir une tablette et un ordinateur à écran tactile connectés à l'Internet et équipés d'une variété de logiciels (suite bureautique, outils de traductique, système de RV, clavier virtuel, etc.), ainsi qu'un stylet.



**Figure 4. Prototype d'un environnement de TDI**

Des données quantitatives et qualitatives ont été relevées au cours de l'expérience, les premières touchant à des aspects tels que la vitesse de frappe au clavier et de lecture à voix haute des traducteurs, l'exactitude de la transcription et l'usage de la souris. Les données qualitatives ont été récoltées lors d'entretiens semi-dirigés avec les traducteurs. L'un des objectifs de ces entretiens était d'extraire les perceptions des traducteurs sur leur utilisation d'un environnement de TDI pour traduire un texte, comparé à leur environnement de travail habituel (ordinateur, clavier et souris). L'hypothèse sous-jacente à l'expérience était que l'environnement de TDI améliorerait l'expérience de travail du traducteur grâce à une interaction plus conviviale avec la machine. Les résultats que nous présentons ici

ne concernent que les aspects ergonomiques évoqués par les traducteurs[5], puisque ces derniers ont été un sujet fort récurrent dans les conversations.

Tout d'abord, dans leur configuration de travail habituelle, les traducteurs sont dotés d'un ordinateur, d'un écran, d'une souris et d'un clavier. Toutefois, pour des questions d'amélioration de la productivité et de prévention de maladies, les configurations de travail de quelques traducteurs ont été modifiées. Par exemple, certains travaillent avec deux écrans qu'ils utilisent respectivement pour afficher le texte source et le texte cible ou pour faire apparaître les deux textes sur un écran et utiliser le second pour la recherche documentaire. Deux participants ont, eux, reçu un clavier et une souris ergonomiques qui leur permettent de placer leurs mains et poignets dans des positions « optimales ». L'un d'entre eux précise aussi avoir reçu l'aide d'un spécialiste en ergonomie pour obtenir une chaise ajustée à sa taille et un placement en angle de ses deux écrans.

Hormis une personne qui se dit entièrement satisfaite de son environnement de travail actuel et indique n'avoir « jamais ressenti [de] fatigue au travail », les autres participants admettent vivre des douleurs physiques, particulièrement aux poignets, dues à la position assise devant un écran qu'ils doivent maintenir pendant de longues heures. Parmi les moyens que déploient les traducteurs pour pallier ces problèmes se trouve la variété des tâches qui les amènent à ne pas toujours faire les mêmes mouvements. Ainsi, deux traducteurs indiquent imprimer le premier jet de leur traduction afin d'appliquer les révisions avec un crayon, ce qui les aiderait aussi à mieux repérer les erreurs. Un autre traducteur choisit aussi d'imprimer son texte source et de l'annoter au crayon de toutes les informations qui lui seront utiles pour traduire. Finalement, le kinésithérapeute d'un des participants lui aurait conseillé de prendre des pauses et de se déplacer régulièrement, mais cela semble difficile à mettre en pratique en raison des délais et des objectifs de productivité journaliers à respecter.

On le voit donc, tous les ajustements sont dus au fait que les traducteurs sont, en grande majorité, soucieux et conscients des implications que la configuration de leur environnement de travail peut avoir sur leur santé. Le participant qui a reçu l'appui du spécialiste en ergonomie dit même avoir voulu recevoir une évaluation dès le début de son activité professionnelle, car il sait que les traducteurs sont sujets aux tendinites du fait qu'ils passent, selon lui, 95 % de leur temps à taper à l'ordinateur. Dans le contexte qui est le sien, il juge son environnement de travail ergonomique et optimal. En revanche, il reconnaît qu'avec un matériel comme celui du prototype d'un environnement de TDI, les traducteurs seraient « dans une classe à part ».

Cela nous amène aux commentaires des participants au sujet de leur utilisation de cet environnement de TDI. Sans grande surprise, les traducteurs ont apprécié la possibilité de travailler tout en pouvant se déplacer et ont évoqué les bienfaits de cette option pour les mêmes aspects de santé qui les préoccupent dans leur contexte de travail actuel. Notons que même le traducteur qui indiquait être entièrement satisfait de son environnement professionnel a reconnu aimer pouvoir bouger en travaillant et a admis qu'il aimerait avoir accès aux deux environnements de travail. L'augmentation de la mobilité apportée par l'interface multimodale de TDI a aussi été soulignée par une personne qui a vu un avantage à pouvoir passer régulièrement d'une interface à une autre et d'un appareil à un autre. Selon elle, cela force aussi le traducteur à bouger du fait qu'il utilise le stylet, les doigts ou la tablette.

Selon un participant, l'ergonomie de l'environnement de TDI pourrait avoir un impact positif sur la qualité de la traduction puisqu'il favoriserait la concentration optimale du traducteur et préserverait son niveau d'énergie au cours de la journée. Ce constat de qualité améliorée du produit de la traduction est partagé par la majorité des traducteurs. Seules deux personnes pensent différemment, l'une ayant démontré un manque d'enthousiasme généralisé face à l'expérience et une autre doutant que la qualité de la traduction puisse être meilleure que celle obtenue à partir d'un environnement de travail traditionnel. Certains participants jugent que l'environnement de TDI diminue le niveau de stress en augmentant le nombre de mots traduits par minute et, du même coup, en accordant plus de temps à la préparation et à la recherche documentaire, ainsi qu'à la révision du texte cible. « Ce sont des choses qu'on n'a presque plus le temps de faire parce que les échéances sont trop serrées », indique l'un d'entre eux. Ce gain de temps se ressentira donc aussi sur la qualité du produit fini.

Toutefois, le gain de temps reste une estimation de la part des traducteurs et il n'a pas été ressenti lors de l'expérience. En effet, la plupart d'entre eux ont signalé que pour tirer réellement profit de l'environnement multimodal, il leur faudrait plusieurs heures de pratique, ne serait-ce que pour entraîner le système de RV à reconnaître les caractéristiques langagières de son utilisateur : « Côté productivité, je pense que ça ferait un creux avant de remonter. Il y a beaucoup de choses à intégrer. » Néanmoins, la majorité ayant pu observer les avantages de la dictée comparée à la saisie au clavier d'ordinateur, les traducteurs de notre étude tendent à voir le potentiel d'un tel environnement de TDI.

## Conclusion

Dans cet article, nous avons tenté de mettre l'accent sur le potentiel ergonomique qu'offrent les technologies de reconnaissance vocale (RV) et les interfaces multimodales pour la traduction professionnelle. Deux études explorant l'usage de la RV et de la traduction dictée interactive (TDI) ont été présentées et nous avons discuté leurs résultats en examinant les avantages de la RV et de la TDI pour l'amélioration et le maintien du bien-être physique et psychologique des traducteurs.

L'ergonomie va bien au-delà de la simple fabrication de chaises de bureau confortables ou de claviers d'ordinateur courbés pour épouser la forme des poignets; c'est pourquoi, depuis longtemps, les études sur le sujet intègrent des questions touchant aux impacts physiques et psychologiques des technologies sur les êtres humains et sur leurs activités. C'est le cas, par exemple, des recherches considérant les effets de l'usage des ordinateurs sur l'acte d'écrire et sur la pensée créative qui y est associée (Daiute 1985). Ces études, souvent initiées suite à des plaintes d'utilisateurs, aident à améliorer les outils, les systèmes ou les logiciels de façon à ce qu'ils répondent aux besoins ergonomiques des personnes qui les emploient. Cette approche permettrait alors d'éviter les nombreux « dérapages » de conception d'outils technologiques, dus à la vision étroite des « experts » qui détiennent une importante formation scientifique ou technique, mais cette dernière étant dépourvue de toute réflexion humaniste (Vicente 2004 : 44-50) :

Les désagréments et les frustrations que nous éprouvons quotidiennement en utilisant nos appareils et nos gadgets sophistiqués ne constituent que la pointe de l'iceberg. Nous devons prendre conscience que ces problèmes appartiennent aussi aux systèmes complexes responsables de notre bien-être et de notre sécurité. (*ibid.* : 39)

Une révolution « techno-humaine » (*ibid.* : 56) s'impose donc et la technologie se doit d'être conçue en meilleure adéquation avec l'être humain qui en fait usage et ses besoins. Dans le cas des traducteurs, il s'agira alors de prendre en considération leur utilisation quotidienne des outils informatiques à leur disposition et les limites qu'ils présentent pour leur proposer une interface qui devienne « une extension naturelle des capacités et habiletés du langagier » (Taravella et Villeneuve 2011, paragr. 51).

En cela, la RV et les interfaces multimodales se présentent comme des solutions ergonomiques prometteuses puisqu'elles fonctionnent déjà très bien au sein de divers domaines professionnels et situations de la vie quotidienne. Leur intégration aux applications de traductique actuelles permettrait alors d'en combler les manques et de prendre en compte le facteur humain. La RV contribue déjà au grand retour de la traduction dictée (TD) dans la pratique de la traduction, qui, selon Gouadec (2007 : 363), est en passe de redevenir la norme :

In all probability, the translation industry in the twenty-first century will go oral. Voice recognition and dictation software is probably going to radically change the way translation is perceived and practiced [...] Dictated translation will become the norm once again.

Cependant, un usage multimodal et ergonomique de la TD se doit de dépasser les pratiques des années 1960 et 1970, période à laquelle il était commun de traduire à voix haute. C'est pourquoi nous envisageons l'intégration de la TDI, à la fois à la formation et à l'environnement de travail des traducteurs. L'interaction multimodale vise à améliorer la relation entre le langagier et la technologie dans tous les aspects de son travail (préparation, production, révision des traductions) pour accroître son efficacité, sa créativité et son bien-être physique et psychologique. En faisant la promotion d'une adaptation optimale de la RV et des technologies interactives émergentes (ex. : écrans tactiles et appareils mobiles), la TDI propose donc d'apporter aux environnements de traductique cette dimension humaine tant attendue.

## Bibliographie

- Barabé, Donald (2013) « Société, technologie et traduction : perspectives et impacts », *The Journal of Specialised Translation* 19: 41–61.
- Benis, Michael (2002) « Softly spoken or hard of hearing? », *Language International* 14, no. 3: 26–9.
- Bowker, Lynne (2002) *Computer-Aided Translation Technology: A Practical Introduction*, Ottawa, University of Ottawa Press.
- Brousseau, Julie, Caroline Drouin, George Foster, Pierre Isabelle, Roland Kuhn, Yves Normandin et Pierre Plamondon (1995) « French speech recognition in an automatic dictation system for translators: The TransTalk project », *Proceedings of Eurospeech '95*, URL : <http://www.iro.umontreal.ca/~foster/papers/ttalk-eurospeech95.pdf> (consulté le 3 mars 2017).
- Brown, P. F., S. F. Chen, S. A. Della Pietra, V. J. Della Pietra, A. S. Kehler et R. L. Mercer (1994) « Automatic speech recognition in machine-aided translation », *Computer Speech and Language* 8, no. 3: 177–87.
- Carl, Michael, Barbara Dragsted et Arnt Lykke Jakobsen (2011) « On the Systematicity of Human Translation Processes », *Actes de la conférence de Tralogy, Paris, le 3 mars 2011*, URL : <http://odel.irevues.inist.fr/tralogy/index.php?id=103> (consulté le 3 mars 2017).
- Ciobanu, Dragoş (2014) « Of Dragons and Speech Recognition Wizards and Apprentices », *Revista Tradumática* 12: 524–38, URL : [https://ddd.uab.cat/pub/tradumatica/tradumatica\\_a2014n12/tradumatica\\_a2014n12p524.pdf](https://ddd.uab.cat/pub/tradumatica/tradumatica_a2014n12/tradumatica_a2014n12p524.pdf) (consulté le 3 mars 2017).
- Ciobanu, Dragoş (2016) « Automatic Speech Recognition in the Professional Translation Process », *Translation Spaces* 5, no.1: 124–44.
- Cohen, Philip, Michael Johnston, David McGee, Sharon L. Oviatt, Josh Clow et Ira Smith (1998) « The Efficiency of Multimodal Interaction: A Case Study », *Proceedings of the ICSLP '98*, URL : <http://web.media.mit.edu/~earroyo/thesis-backup/Thesis-prep/multimodal-interfaces/Alredy-printed/the-efficiency-of-multimodal-interaction.pdf> (consulté le 3 mars 2017).
- Daiute, Colette (1985) *Writing and Computers*, Reading [MA], Addison-Wesley.
- Désilets, Alain, Marta Stojanovic, Jean-François Lapointe, Richard Rose et Aarthi Reddy (2008) « Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation », *Proceedings of IWSLT 2008*, URL : <http://www.mt-archive.info/IWSLT-2008-Desilets.pdf> (consulté le 3 mars 2017).
- Dragsted, Barbara et Inge Gorm Hansen (2009) « Exploring Translation and Interpreting Hybrids. The Case of Sight Translation », *Meta* 54, no. 3: 588–604, URL : <http://www.erudit.org/revue/meta/2009/v54/n3/038317ar.pdf> (consulté le 3 mars 2017).
- Dragsted, Barbara, Inge Gorm Hansen et Henrik Selsøe Sørensen (2009) « Experts Exposed », *Copenhaguen Studies in Language* 38: 293–317.
- Dymetman, Marc, Julie Brousseau, George Foster, Pierre Isabelle, Yves Normandin et Pierre Plamondon (1994) « Towards an Automatic Dictation System for Translators: the TransTalk Project », *Proceedings of the ICSLP '94*, URL : <http://arxiv.org/abs/cmp-lg/9409012> (consulté le 8 mars 2017).
- Garcia-Martinez, Mercedes, Karan Singla, Aniruddha Tammewar, Bartolomé Mesa-Lao, Ankita Thakur, Anusuya M. A., Michael Carl et Srinivas Bangalore (2014) « SEECAT: ASR & Eye-tracking Enabled Computer-Assisted Translation », *Proceedings of the 17th Annual Conference of the European Association for Machine Translation*: 81–8, URL : <http://www.mt-archive.info/10/EAMT-2014-Garcia-Martinez.pdf> (consulté le 8 mars 2017).
- Girol, David, Ramón Lopez-Cozar, Zoraida Callejas, Gonzalo Espejo et Nieves Ábalos (2012) « On the Development of Adaptive and User-Centered Interactive Multimodal Interfaces », dans *Speech, Image and Language Processing for Human Computer Interaction: Multi-Modal Advancements*, Uma Shanker Tiwary et Tanveer J. Siddiqui (dir.), Hershey [PA], Information Science Reference: 262–91.
- Gouadec, Daniel (2007) *Translation as a profession*, Amsterdam, John Benjamins.
- Jiménez Ivars, Amparo et Amparo Hurtado Albir (2003) « Variedades de traducción a la vista. Definición y clasificación », *TRANS: Revista de Traductología* 7: 47–57, URL : [http://www.trans.uma.es/trans\\_07.html](http://www.trans.uma.es/trans_07.html) (consulté le 3 mars 2017).
- Jurafsky, Daniel et James H. Martin (2009) *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2e éd.), Upper Saddle River [NJ], Pearson Prentice Hall.
- Karamanis, Nikiforos, Saturnino Luz et Gavin Doherty (2011) « Translation practice in the workplace: A contextual analysis and implications for machine translation », *Machine Translation* 25, no. 1: 35–52.

- Lippmann, Richard P (1997) « Speech recognition by machines and humans », *Speech Communication* 22, no. 1: 1–15, URL : [http://doi.org/10.1016/S0167-6393\(97\)00021-6](http://doi.org/10.1016/S0167-6393(97)00021-6) (consulté le 3 mars 2017).
- Mees, Inger M., Barbara Dragsted, Inge Gorm Hansen et Arnt Lykke Jakobsen (2013) « Sound effects in translation », *Target* 25, no. 1: 140–54, URL : <http://openurl.ingenta.com/content/xref?genre=article&issn=0924-1884&volume=25&issue=1&spage=140> (consulté le 3 mars 2017).
- Mesa-Lao, Bartolomé (2014) « Speech-Enabled Computer-Aided Translation: A Satisfaction Survey with Post-Editor Trainees », *Workshop on Humans and Computer-assisted Translation*: 99–103, URL : <http://www.aclweb.org/anthology/W14-0315> (consulté le 8 mars 2017).
- O'Brien, Sharon (2012) « Translation as human-computer interaction », *Translation Spaces* 1, no. 1: 101–22, URL : <http://doi.org/10.1075/ts.1.05obr> (consulté le 8 mars 2017).
- Office québécois de la langue française (2007) *Grand dictionnaire terminologique*, URL : <http://www.granddictionnaire.com> (consulté le 3 mars 2017).
- Oviatt, Sharon (2012) « Multimodal Interfaces », dans *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (3<sup>e</sup> éd), Julie A. Jacko (dir.), New York, Lawrence Erlbaum Associates: 405–29.
- Oviatt, Sharon et Philip R. Cohen (2000) « Multimodal systems that process what comes naturally », *Communications of the ACM* 43, no. 3: 45–53.
- Oviatt, Sharon, Philip Cohen et Michelle Wang (1994) « Toward interface design for human language technology: Modality and structure as determinants of linguistic complexity », *Speech Communication* 15: 283–300.
- Oviatt, Sharon, Rachel Coulston et Rebecca Lunsford (2004) « When do we interact multimodally? Cognitive load and multimodal communication partners », *Proceedings of the ICMI'04*, New York, ACM Press: 129–36.
- Oviatt, Sharon, Rebecca Lunsford et Rachel Coulston (2005) « Individual differences in multimodal integration patterns: What are they and why do they exist? », *CHI'05 Proceedings*, New York, ACM Press: 241–9.
- Oviatt, Sharon et Robert van Gent (1996) « Error resolution during multimodal human-computer interaction », *Proceedings of the ICSLP'96*: 204–7.
- Rabiner, Lawrence (1997) « Applications of speech recognition in the area of telecommunications », *Proceedings of the IEEE'97*, URL : [http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/341\\_telecom\\_applications.pdf](http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/341_telecom_applications.pdf) (consulté le 3 mars 2017).
- Reddy, Aarthi et Richard Rose (2010) « Integration of Statistical Models for Dictation of Document Translations in a Machine Aided Human Translation Task », *IEEE Transactions on Audio, Speech and Language Processing* 18, no. 8: 1–11, URL : <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05393062> (consulté le 3 mars 2017).
- Romero-Fresco, Pablo (2011) *Subtitling Through Speech Recognition: Respeaking*, Manchester, St. Jerome.
- Schalkwyk, Johan, Doug Beeferman, Bill Byrne, Ciprian Chelba, Mike Cohen, Maryam Garret et Brian Strope (2010) « Google Search by Voice: A Case Study », *Advances in Speech Recognition*, New York, Springer: 61–90, URL : [http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/en/us/pubs/archive/36340.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/pubs/archive/36340.pdf) (consulté le 3 mars 2017).
- Seaman, Lee (2002) « Voice Recognition for Translators, Or Why I Started Talking To My Computer », *JLD Times. Newsletter of the Japanese Language Division of the American Translators Association*: 6–8.
- Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3<sup>rd</sup> ed.). Reading, Massachusetts : Addison-Wesley.
- Stroman, John (2002) « Translation And Voice Recognition Software », *JLD Times. Newsletter of the Japanese Language Division of the American Translators Association*: 1–5.
- Taravella, Anne-Marie et Alain Villeneuve (2011) « Aspects humains des technologies langagières dans l'organisation », *Actes de la conférence de Tralogy*, Paris, le 3 mars 2011, URL : <http://lodel.irevues.inist.fr/tralogy/index.php?id=134> (consulté le 3 mars 2017).
- Taravella, Anne-Marie et Alain Villeneuve (2013) « Acknowledging the needs of computer-assisted translation tools users: the human perspective in human-machine translation », *The Journal of Specialised Translation* 19: 62–74, URL : [http://www.jostrans.org/issue19/art\\_taravella.pdf](http://www.jostrans.org/issue19/art_taravella.pdf) (consulté le 3 mars 2017).
- Torres-Hostench, Olga, Joss Moorkens, Sharon O'Brien et Joris Vreeke (2017). Testing interaction with a Mobile MT post-editing app. *Translation & Interpreting* 9, no. 2:138-150, URL: <http://www.trans-int.org/index.php/transint/article/view/645/291> (consulté le 8 aout 2017).
- Vicente, Kim (2004) *Le facteur humain : réinventer notre rapport à la technologie*, trad. N. Guillet, Outremont, Logiques.
- Vidal, Enrique, Francisco Casacuberta, Luis Rodríguez, Jorge Civera et Carlos D. Martínez Hinarejos (2006) « Computer-assisted translation using speech recognition », *IEEE Transactions on Audio, Speech and Language Processing* 14, no. 3, URL : <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01621206> (consulté le 3 mars 2017).
- Williamson, Julie R., Andrew Crossan et Stephen Brewster (2011) « Multimodal Mobile Interactions: Usability Studies in Real World Settings », *Proceedings of the 13th International Conference on Multimodal Interaction*: 361–8.
- Wu, Ming (2008) *The Most Cost-effective Way of Translation: Combined Use of Digital Recorders and Dragon Voice Recognition System*, Diaporama d'une présentation aux Nations-Unies à Nairobi, le 4 aout. Inédit.
- Zapata, Julián (2012) *Traduction dictée interactive : intégrer la reconnaissance vocale à l'enseignement et à la pratique de la traduction professionnelle*, Mémoire de maîtrise, Université d'Ottawa, URL : [http://www.ruor.uottawa.ca/en/bitstream/handle/10393/23227/Zapata\\_Rojas\\_Julian\\_2012\\_these.pdf?sequence=1](http://www.ruor.uottawa.ca/en/bitstream/handle/10393/23227/Zapata_Rojas_Julian_2012_these.pdf?sequence=1) (consulté le 3 mars 2017).
- Zapata, Julián (2016) *Translators in the Loop: Observing and Analyzing the Translator Experience with Multimodal Interfaces for Interactive Translation Dictation Environment Design*, Thèse de doctorat, University of Ottawa,

URL : [https://www.ruor.uottawa.ca/bitstream/10393/34978/1/Zapata\\_Rojas\\_Julian\\_2016\\_thesis.pdf](https://www.ruor.uottawa.ca/bitstream/10393/34978/1/Zapata_Rojas_Julian_2016_thesis.pdf) (consulté le 3 mars 2017).

Zapata, Julián et Andreas S. Kirkedal (2015) « Assessing the Performance of Automatic Speech Recognition Systems When Used by Native and Non-Native Speakers of Three Major Languages in Dictation Workflows », *Proceedings of the 20th Nordic Conference of Computational Linguistics*: 201–10, URL : <http://www.aclweb.org/anthology/W15-1825> (consulté le 3 mars 2017).

Zapata, Julián et Jean Quirion (2016) « La traduction dictée interactive et sa nécessaire intégration à la formation des traducteurs », *Babel* 62, no. 4: 531–51, URL : <https://benjamins.com/#catalog/journals/babel.62.4.01zap/details> (consulté le 3 mars 2017).

Zapata, Julián, Sheila Castillo et Joss Moorkens (2017) « Translation Dictation vs. Post-editing with Cloud-based Voice Recognition: A Pilot Experiment », *Proceedings of The Sixteenth MT Summit Conference*, Nagoya, Japon, 18-22 septembre, 2017.

## Notes

[1] Dans le présent article, le genre non marqué, c'est-à-dire le masculin, est utilisé uniquement pour alléger le texte et il inclut le genre féminin là où il y a lieu.

[2] L'ergonomie est « l'étude scientifique des conditions (psychophysiologiques et socioéconomiques) de travail, de l'adaptation des outils, postes de travail aux utilisateurs, des relations entre l'utilisateur et la machine » (*Petit Robert de la langue française en ligne*, 2017).

[3] Nous reconnaissons qu'il existe une terminologie variée pour désigner la notion que nous choisissons d'appeler « reconnaissance vocale ». En français, certains chercheurs s'y référeront à l'aide de termes tels que « reconnaissance automatique de la parole », « reconnaissance de la voix » ou simplement en employant l'acronyme anglais « ASR » (*Automatic Speech Recognition*). Pour notre part, nous nous appuyons sur la terminologie et la définition proposées par l'Office québécois de la langue française.

[4] Selon une enquête menée par Verástegui en 2009 auprès des participants de la Réunion internationale annuelle sur la traduction et la terminologie assistées par ordinateur (JIAMCATT) regroupant des représentants de grands services de traduction. Les résultats de l'enquête ont été distribués par l'auteur via LISTSERV aux participants de JIAMCATT 2009, mais n'ont fait l'objet d'aucune publication.

[5] Pour en savoir plus sur les autres résultats de l'expérience, nous renvoyons le lecteur à Zapata (2016).

---

©inTRAlinea & Julián Zapata & Elizabeth C. Saint (2017).

"Traduire à voix haute: la traduction dictée interactive comme solution ergonomique", *inTRAlinea* Vol. 19.

Stable URL: <http://www.intraline.org/archive/article/2261>