



HAL
open science

L'analyse en temps réel de la production du langage écrit

Thierry Olive

► To cite this version:

Thierry Olive. L'analyse en temps réel de la production du langage écrit. Clara Solier; Lucille Soulier; Nour Ezzedine. Introduction aux statistiques en sciences du langage. Traitement et analyse de données avec R, Dunod, pp.69-94, 2023, Univers Psy, 978-2-10-085264-2. <hal-04369284>

HAL Id: hal-04369284

<https://hal.science/hal-04369284v1>

Submitted on 3 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

Chapitre 3

L'analyse en temps réel de la production du langage écrit

Thierry Olive, Directeur de recherche au CNRS

Centre de Recherches sur la Cognition et l'Apprentissage (CNRS & Université de Poitiers, UMR 7295)

Thierry.olive@univ-poitiers.fr

Résumé. L'étude en temps réel de la production du langage écrit vise à décrire le fonctionnement cognitif des rédacteurs et la dynamique de la production écrite. Pour cela, l'activité d'écriture est enregistrée et diverses variables informant sur le processus d'écriture sont collectées et analysées. Ce chapitre présente un panorama des principales méthodes et techniques qui peuvent être utilisées pour, d'une part, étudier les processus cognitifs impliqués dans la production du langage écrit, du tracé de lettres à la production de textes et d'autre part pour analyser la dynamique du texte « en train de se faire » (la textualisation, Cislaru & Olive, 2018).

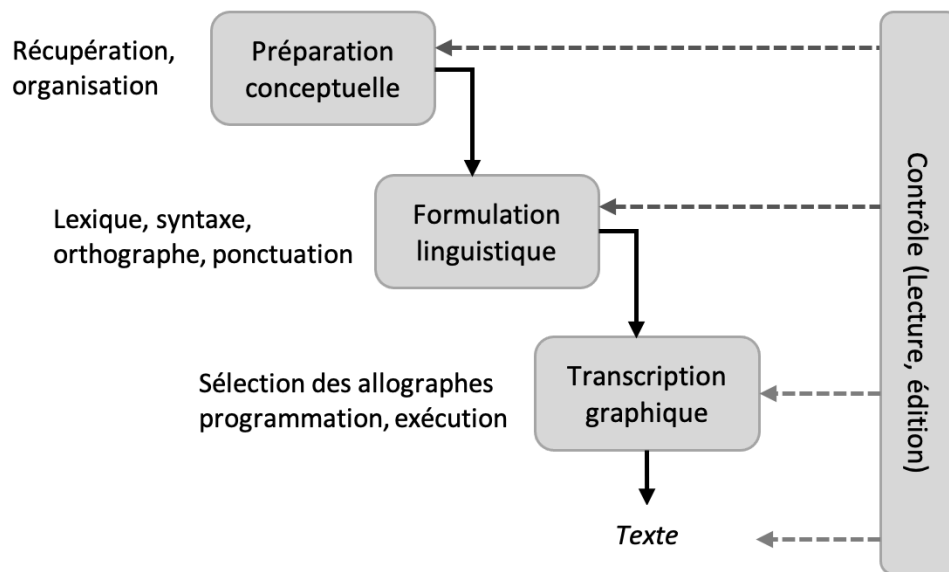
1. L'étude en temps réel de la production écrite

Étudier la production écrite nécessite au moins deux approches : la première, centrée sur le produit, analyse le langage produit. Cette analyse peut porter sur différentes dimensions comme l'orthographe, la richesse lexicale, la structure syntaxique, la cohésion, la cohérence, le schéma textuel, etc. Il s'agit d'interpréter les variations de contenu et de langage en fonction de variations, par exemple contextuelles. La seconde approche, dite en temps réel, est centrée sur le processus de production. Elle analyse le fonctionnement cognitif des rédacteurs et la dynamique de la production écrite à partir d'enregistrements de l'écriture. Des données comme la durée des pauses ou les vitesses d'écritures sont analysées pour mettre à jour les stratégies cognitives des rédacteurs et le processus de textualisation (Cislaru & Olive, 2018). Ce chapitre traite spécifiquement de l'approche en temps réel. Il présente un panorama des principales méthodes et techniques qui peuvent être utilisées pour étudier le processus de production du langage écrit, du tracé de lettres à la production de textes, en contexte de laboratoire ou dans des cadres plus écologiques.

Avant de présenter en détail ces techniques, il est important de décrire les opérations cognitives mises en œuvre lors de la production écrite. Quatre composantes cognitives sont impliquées pour produire du langage écrit : une première concerne les traitements sémantiques, dits de préparation conceptuelle. Ces traitements sont mis en œuvre pour récupérer les connaissances stockées en mémoire à long terme du rédacteur, rechercher des informations dans des sources externes (des ouvrages, sur internet...), les organiser, et les hiérarchiser. Une composante, dite de formulation, intervient à un niveau linguistique pour transformer le message conceptuel en un message verbal en définissant les propriétés orthographiques, syntaxiques et morphologiques des unités linguistiques traitées. La composante de contrôle ou de révision évalue le produit de chacune des composantes précédentes. La révision peut être mentale et porter sur une partie du message qui n'est pas encore tracée ou intervenir sur le texte déjà écrit. Enfin, une dernière composante, psychomotrice, assure la transcription du texte (cf. Figure 3.1). Un objectif principal des recherches consiste à pister la mise en œuvre de ces processus pour comprendre les stratégies d'écriture des rédacteurs, qu'il s'agisse par exemple de trouver l'orthographe d'un mot peu familier ou de rédiger le récit d'un événement.

Figure 3.1

Les processus cognitifs engagés pour rédiger un texte (selon Olive, 2014).



Les traitements cognitifs mis en jeu pour récupérer l'information relative aux mots et à leur orthographe impliquent deux voies de traitement (pour une revue, voir Perret & Olive, 2019). La première, sub-lexicale ou de conversion phonème-graphème, est une procédure séquentielle basée sur les fréquences d'association entre phonèmes et graphèmes. La seconde, lexicale, récupère les représentations lexicales/orthographiques stockées en mémoire à long terme. Cette voie lexicale peut elle-même être subdivisée en deux : la représentation phonologique peut être directement appliquée à une représentation orthographique dans un lexique orthographique ou indirectement via une représentation sémantique qui joue le rôle de médiateur entre les représentations phonologiques et orthographiques. La représentation orthographique qui résulte de ces processus est ensuite stockée dans une mémoire de travail orthographique pendant la réalisation de la trace graphique. Ces deux voies fonctionnent en parallèle mais à des vitesses différentes. La voie lexicale dépend des caractéristiques lexicales (par exemple, la fréquence des mots) tandis que la voie sub-lexicale est influencée par des caractéristiques sub-lexicales (par exemple, la fréquence des associations phonémiques de graphèmes). Lorsque la représentation lexicale est générée (ou récupérée), un ensemble de processus traduit les graphèmes sélectionnés en commandes motrices directement utilisables par le système neuromusculaire pour réaliser le tracé graphique.

Un autre champ de recherche concerne l'étude de l'écriture proprement dite, c'est-à-dire d'inscription de la trace écrite. L'écriture manuscrite nécessite trois étapes pour déterminer les caractéristiques des allographes sélectionnés (par exemple, majuscules, minuscules, écriture cursive), le nombre de traits formant chaque allographe, et l'exécution de la séquence de mouvements (Ellis, 1988 ; van Galen, 1991). La frappe au clavier nécessite, quant à elle, la programmation de mouvements plus simples pour atteindre et presser les touches appropriées. Cependant, elle demande une coordination bimanuelle et de plusieurs doigts. Les dactylographes novices encodent en série chaque lettre du mot : ils encodent une première lettre, trouvent la bonne touche, déplacent le doigt vers cette touche, appuient sur la touche et encodent la lettre suivante (Yamaguchi & Logan, 2014). Les dactylographes experts encodent un mot en activant en parallèle plusieurs frappes (Crump & Logan, 2010). Chez ces dactylographes, le contexte des mots active des séquences de frappes qui sont récupérées à la volée en mémoire à long terme (Logan, 2018), comme en témoigne le fait que les intervalles entre les frappes chez

les dactylographes experts sont sensibles à la fréquence d'occurrence des combinaisons de lettres concernées.

2. Les méthodes d'étude en temps réel

Les méthodes en temps réel analysent les activités cognitives en même temps qu'elles se produisent. Elles décrivent le décours temporel des processus de production et leur agencement. Pour cela, des données susceptibles de rendre compte de l'activation des processus mis en œuvre sont recueillies pendant l'activité à partir d'un enregistrement de l'écriture. Ces méthodes reposent sur la chronométrie mentale dont elles exploitent deux des postulats : les processus psychologiques ont une durée qui peut être mesurée et les changements de durée reflètent des changements dans la nature ou la complexité du traitement requis (Cattell, 1886).

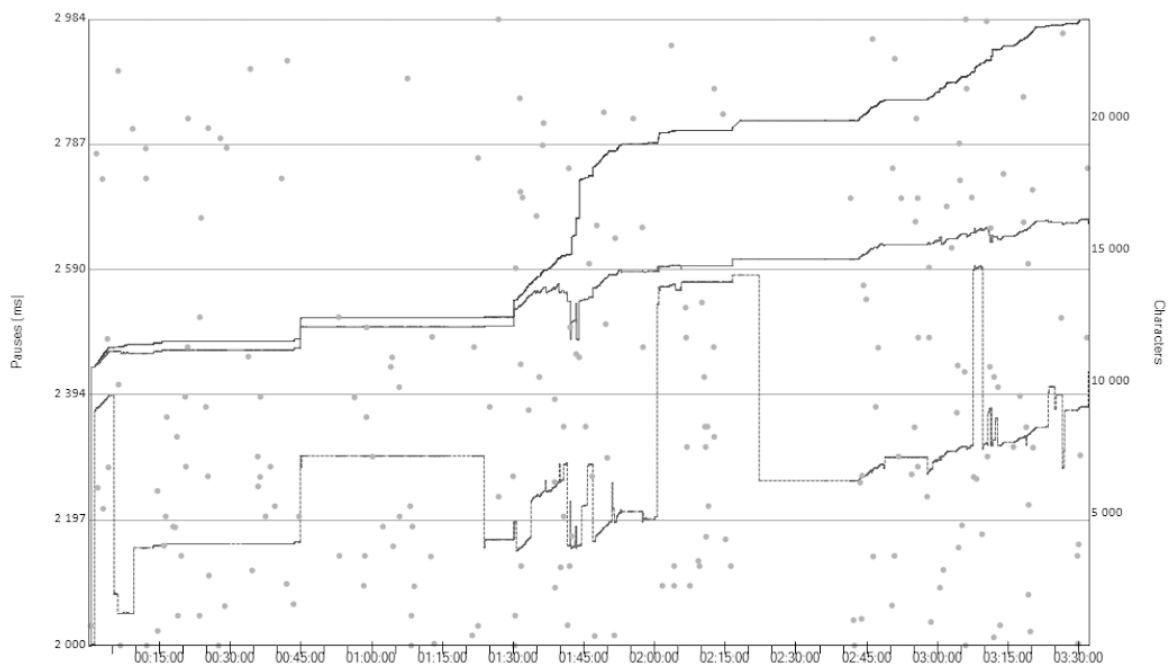
2.1. L'enregistrement de l'écriture

L'enregistrement de l'écriture manuscrite est piloté par un ordinateur. Un logiciel enregistre l'activité manuscrite d'un rédacteur qui écrit sur une feuille posée sur une tablette graphique avec un stylo électronique. Les données spatio-temporelles relatives au tracé sont enregistrées, généralement sous la forme d'une suite de coordonnées verticales et horizontales (x, y) auxquelles est associée la chronologie des événements. Ces logiciels peuvent restituer ces informations en faisant apparaître sur l'écran le tracé de l'écriture. Ils peuvent aussi proposer des pré-analyses descriptives des données, comme la durée de pause moyenne, le nombre de pauses, etc. Par exemple, le logiciel *Eye and Pen* enregistre simultanément la trace graphique manuscrite et l'activité oculaire des rédacteurs (pour une présentation rapide de la dernière version du logiciel, voir Chesnet et al., 2022). Une version récente permet également de synchroniser les enregistrements de l'écriture avec l'activité électro-encéphalographique du rédacteur. Il permet aussi de visualiser les pauses dans le texte ou les zones de regard des rédacteurs. Un autre logiciel, *Handspy*, développé par Alves, Leal et Limpo (2019), enregistre l'écriture au moyen d'un stylo électronique : le rédacteur écrit sur une feuille de papier spécifique qui permet à une caméra située à la pointe du stylo de localiser la trace écrite sur la feuille.

Les logiciels dits d'enregistrement de frappes, ou de *keystroke logging* en anglais, enregistrent l'ensemble des gestes de production écrite sur clavier. Les frappes sur le clavier (caractères, espaces, effacements), les mouvements de souris pour replacer le curseur dans le texte ou pour sélectionner des éléments sont enregistrés avec les données temporelles associées. La production du texte peut ensuite être visionnée grâce à une fonction de relecture et des préanalyses des données de pause et de vitesse sont proposées. Ces préanalyses sont souvent plus nombreuses et variées avec la frappe au clavier, les logiciels pouvant automatiquement isoler différentes unités de langage dans les textes (mots, phrases...). Plusieurs solutions d'enregistrements dynamiques sont accessibles. Par exemple, *Inputlog* (Leijten & van Waes 2006) enregistre toutes les frappes au clavier et les actions de la souris et propose des statistiques de base sur les pauses et les opérations de révisions (voir figure 3.2). *Scriptlog* (Strömquist & Malmsten, 1998) intègre dans un programme unique un logiciel simplifié de traitement de texte ainsi que des outils d'enregistrement et d'analyse. Enfin, ces logiciels peuvent être couplés à des dispositifs d'enregistrement des mouvements oculaires des rédacteurs (pour plus de détails sur les types d'analyse développés dans ce cadre, voir Wengelin et al., 2009).

Figure 3.2

Représentation de l'activité d'un rédacteur proposée par le logiciel d'enregistrement de l'écriture au clavier Inputlog.



Note : Les points indiquent des pauses dont la durée (en millisecondes) est indiquée sur l'axe vertical gauche et le moment de survenue est indiqué sur l'axe horizontal. La courbe supérieure indique la somme cumulée de caractères frappés depuis le début de la session de rédaction. La courbe médiane représente la somme des caractères présents dans le texte : une courbe descendante indique donc une révision et la suppression de caractères. La courbe inférieure, pointillée, indique la position du curseur dans le texte et peut signaler des retours dans le texte.

L'écriture enregistrée est utilisée en psychologie cognitive, au travers d'études expérimentales, mais aussi dans des approches plus écologiques comme en génétique textuelle pour analyser les opérations de révision ou encore dans des approches plus linguistiques pour étudier par exemple les jets textuels, les segments de textes produits de façon ininterrompue. Enfin, selon les questions posées, différentes variables sont analysées et des logiciels plus spécifiques peuvent être utilisés. Ainsi, les travaux sur la production de mots exploitent les données relatives à la latence de production et à la fluence d'écriture des mots (pause, vitesse, pics d'accélération, etc.). Des logiciels qui facilitent la présentation des stimuli tout en enregistrant l'écriture sont plus fréquemment utilisés dans ce champ (par exemple DMDX, Forster & Forster, 2003). La recherche sur la production de textes analyse essentiellement les pauses de production, et plus précisément leur durée, leur fréquence, leur localisation dans le texte, etc. La partie suivante décrit les méthodes utilisées dans les recherches sur la production de mots. Les travaux sur la production de textes sont décrits dans un second temps.

2.2. Étudier la production de mots : les tâches de dénomination, de dictée et de copie

La recherche sur la production écrite de mots isolés peut s'appuyer sur l'analyse des erreurs spontanées de locuteurs typiques (e.g., Berg, 2002) ou de patients atteints de lésions cérébrales (e.g., Caramazza & Miceli, 1990). L'observation des erreurs est à l'origine des premières propositions théoriques sur les processus cognitifs impliqués dans la production écrite des mots même si l'inférence à partir des observations neuropsychologiques peut être limitée. Les erreurs observées peuvent en effet résulter de processus affectés et ne pas apparaître avec un fonctionnement typique. De plus, ces analyses sont limitées : par exemple, le nombre et la

nature des erreurs recueillies et analysées ne sont pas contrôlables et dépendent du matériel recueilli, ce qui peut en partie biaiser la validité du corpus et donc les observations et interprétations ultérieures. En outre, les observations neuropsychologiques n'informent pas sur la dynamique des processus. Aussi, le recours à des méthodes en temps réel est devenu la norme dans les travaux sur la production de mots. C'est le cas des méthodes de dénomination, de dictée et de copie de mots.

Dans la version classique de la dénomination d'images, le scripteur voit une image et il doit écrire le plus rapidement possible le mot qui lui est associé. L'hypothèse sous-jacente est que la dénomination d'images mobilise les mêmes processus que ceux qui sont impliqués dans la production de mots basée sur les concepts. Aussi, le temps écoulé entre la présentation du stimulus déclenchant la production de mots et le début de l'écriture – la latence – devrait fournir des informations sur le type de représentation activé et le déroulement temporel de leur traitement. La dénomination d'images implique une première étape conceptuelle pour traiter l'image. Ces traitements opèrent à un niveau structural (afin de permettre une reconnaissance perceptive de l'objet) et à un niveau sémantique (les connaissances du rédacteur sur l'objet présenté). Les informations sémantiques activent des représentations lexicales orthographiques et phonologiques en parallèle, via les deux voies d'encodage orthographique.

En production sous dictée, le traitement débute par une analyse acoustique du stimulus sonore. L'analyse des traits phonétiques permet de reconnaître les phonèmes et d'activer sur leur base, une entrée lexicale correspondante. Deux voies de traitement sont disponibles. La première, la voie lexicale (orthographique), active à partir du lexème phonologique le système sémantique. Les représentations sémantiques permettent la sélection d'un lexème au sein du lexique orthographique. La voie phonologique débute à partir des phonèmes reconnus durant le traitement de l'entrée. La seconde voie, phonologique sous-lexicale, débute par la reconnaissance des phonèmes et poursuit par la conversion de chaque phonème en graphème.

Les tâches de dénomination et de dictée impliquent donc deux étapes de traitement orthographiques : une première analyse l'entrée pour une reconnaissance, une deuxième récupère les codes orthographiques. Toutefois, la voie orthographique lexicale est plus fortement mise en jeu durant une tâche de production à partir d'images (Bonin et al., 2015). À l'inverse, en tâche de production sous dictée, où l'entrée dans le système cognitif est de nature phonologique, la voie phonologique est plus impliquée. Une troisième, motrice, convertit les représentations orthographiques en représentations motrices utilisables par le système moteur.

Dans les tâches de copie, les scripteurs lisent des mots qu'ils doivent copier dès qu'ils sont présentés et la latence et la vitesse de copie peuvent être mesurées (pour une revue, Lambert et al., 2011). La copie minimise l'intervention des traitements conceptuels et sémantiques. Lorsque le mot est copié, il peut l'être une seule fois ou plusieurs fois de suite et l'item à copier ne peut être relu après la première identification. Les mesures ne portant pas sur la première copie mettant en jeu des processus de reconnaissance visuelle sont minimisés. De ce fait, les processus ne concernent potentiellement que le lexique orthographique, le buffer graphémique, ainsi que les processus graphomoteurs. Les latences entre deux copies du même mot sont donc supposées ne refléter que des processus orthographiques et/ou graphomoteurs.

Les paramètres de ces différentes tâches peuvent être modifiés pour créer des interférences ou pour amorcer des traitements cognitifs (e.g., Bonin & Fayol, 2002). Par exemple, le stimulus cible peut être associé à un distracteur et la tâche du participant est de nommer le stimulus cible tout en ignorant le distracteur (visuel ou auditif). Le distracteur peut soit amorcer soit interférer avec la dénomination, créant respectivement un effet de facilitation ou d'interférence sur les latences. Si le temps mis pour dénommer ou écrire le nom de la cible augmente lorsque le distracteur est relié à la cible, on parle d'effet inhibiteur de la cible. Au contraire, si le temps de traitement de la cible diminue, on parle d'effet facilitateur. Il est également possible de manipuler l'intervalle entre la présentation du distracteur et celle du stimulus cible, le *Stimulus*

Onset Asynchrony (SOA), afin d'examiner le moment où les représentations mentales sont activées. Le SOA est négatif lorsque le distracteur est présenté avant la cible et il est positif lorsque le distracteur succède la cible. De plus, la durée de présentation de l'amorce et sa visibilité peuvent varier. Par exemple, sa durée de présentation peut être inférieure au seuil de perception ou supérieure.

Si la latence de production est la variable qui a reçu le plus d'attention, les indices de fluidité de l'écriture traduisent l'aisance des traitements cognitifs. La fluidité peut être évaluée en mesurant le temps entre la production de chaque lettre qui compose un mot. Cependant, le style d'écriture des participants peut poser des problèmes : avec une écriture cursive, très peu ou pas de frontières apparaissent entre les lettres. Il est alors possible de demander aux participants d'utiliser une écriture non cursive mais scripte afin de distinguer les frontières des lettres. Kandel et al. (2006) ont par exemple demandé aux participants d'utiliser une écriture majuscule. La généralisation des résultats obtenus avec une écriture peu familière peut néanmoins être limitée. Les études qui font appel à la dactylographie plutôt qu'à l'écriture manuscrite sont plus propices aux analyses des intervalles inter-lettres car une pause apparaît entre chaque frappe ou action au clavier, mais elles imposent de prendre en compte la variabilité des compétences dactylographiques, qui est plus variable que l'écriture manuscrite automatisée des rédacteurs adultes. La fluidité de l'écriture peut également être évaluée en mesurant le nombre de micro-accélération, ou pics de vitesse, qui se produisent lors de la réalisation d'une trace écrite (e.g., Kandel & Perret, 2015). Cependant, la mesure des pics de vitesse se heurte aux mêmes écueils que la durée d'écriture : compter le nombre de pics de vitesse qui se produisent lors de la production d'une unité écrite nécessite d'en définir empiriquement les limites. La durée totale de l'écriture des mots peut aussi être utilisée. Cependant, comme les mots diffèrent en longueur, leur longueur doit être prise en compte, par exemple en divisant la durée d'un mot par le nombre de ses lettres. Cela ne résout donc pas complètement le problème : une lettre nécessite un certain nombre de traits qu'il faut tracer pour la produire et le nombre de traits entre lettres n'est pas constant. Une solution peut donc consister à diviser la durée du mot par le nombre de traits qui le composent (Séraphin-Thibon et al., 2018). Au-delà des difficultés à définir le nombre de traits, l'analyse des traits de l'écriture cursive est difficile à réaliser car il est nécessaire de définir le début et la fin d'une lettre afin de mesurer sa durée.

2.3. Étudier la production de textes

Les tâches utilisées dans les recherches sur la production de textes reposent généralement sur la lecture d'une consigne invitant le rédacteur à produire un texte structuré et de qualité. Selon les études, ces consignes manipulent des dimensions contextuelles, individuelles... Les tâches de production de textes engagent généralement tous les processus rédactionnels, même si certaines peuvent minimiser l'intervention de certains traitements. C'est le cas par exemple des tâches de copie qui minimisent l'intervention des processus de conceptualisation, de formulation et de révision. L'objectif principal des études en temps réel de la production de textes est de comprendre le fonctionnement cognitif du rédacteur, c'est-à-dire la façon dont les processus rédactionnels sont coordonnés.

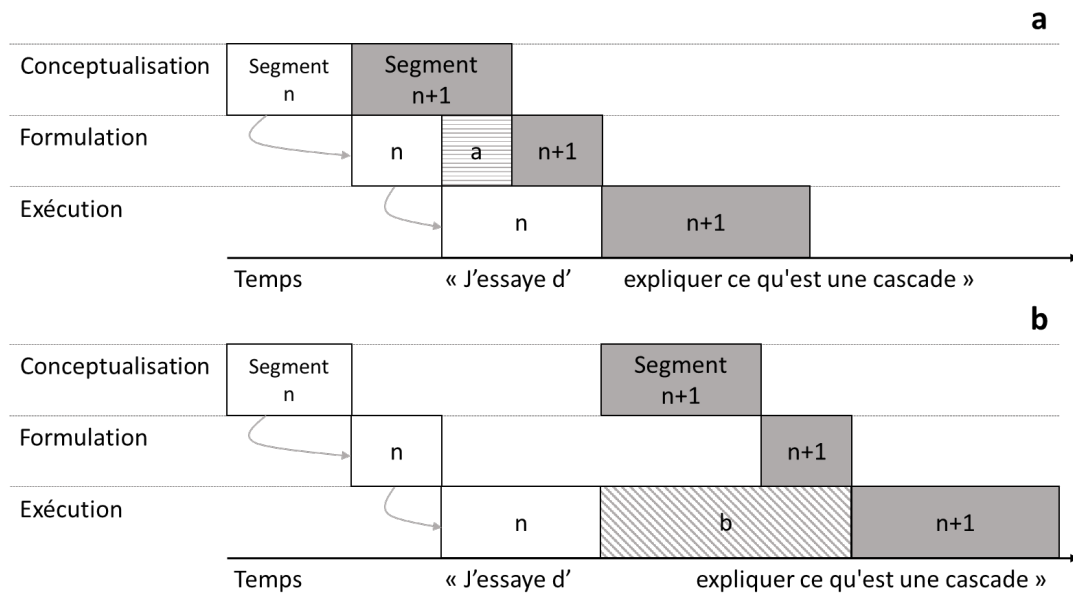
La plupart des processus rédactionnels imposent de fortes demandes cognitives aux rédacteurs et produire un texte sollicite donc fortement leur mémoire de travail, qui a une capacité limitée (Olive, 2022). Leur mise en œuvre s'effectue donc dans le cadre d'une concurrence pour un accès à des ressources cognitives limitées qui doivent être partagées.

La coordination des processus rédactionnels peut reposer sur un fonctionnement séquentiel (*thinking-then-writing*) ou plus parallèle (*thinking-while-writing*) (Olive, 2014). Par exemple, un rédacteur expert rédige un texte avec une stratégie essentiellement parallèle (cf. Figure 3.3,

partie a), mais s'il est en situation de surcharge, il peut opérer avec une stratégie plus séquentielle (cf. Figure 3.3, partie b).

Figure 3.3.

La coordination des processus rédactionnels : (a) mode parallèle (thinking-while-writing) ; (b) mode séquentiel (thinking-then-writing), d'après Olive (2014).



Plusieurs résultats ont confirmé que cette coordination parallèle est possible par exemple lorsque la transcription est automatisée (Olive & Kellogg, 2002). Alamargot et al. (2007) ont observé que les rédacteurs peuvent lire leur texte pour le planifier ou le réviser en même temps qu'ils transcrivent un segment. Chanquoy et al. (1990) ont demandé à des adultes et à des enfants de rédiger des phrases contenant deux propositions. Ils ont enregistré la durée d'écriture de chacune des propositions et celle de la pause inter-propositionnelles. Les adultes rédigeaient leurs phrases avec des durées de pause inter-propositions plus courtes que les enfants. De plus, les adultes accéléraient l'écriture de la fin de la première proposition et de la deuxième proposition. Ces résultats semblent indiquer que les rédacteurs adultes, contrairement aux novices, commencent à planifier la fin de leur phrase durant la rédaction de son début. Ce pattern de résultat a également été montré en production de mots : les processus orthographiques d'un mot sont activés en même temps que leur transcription (Delattre et al., 2006). Sausset et al. (2013) ont observé, avec des mots bi-syllabiques, une augmentation de la durée d'écriture, de la trajectoire ainsi que de la fluence de la première lettre de la deuxième syllabe par rapport à la dernière lettre de la première syllabe. Ce résultat a été interprété comme le reflet d'un traitement parallèle, le rédacteur anticipant la deuxième syllabe pendant la transcription de la première. De telles anticipations ont également été montrées au niveau lexical lors de la production d'un texte (Maggio et al., 2011).

3. L'analyse de la fluence : pauses et débits

3.1. Aspects méthodologiques

L'analyse des pauses de rédaction est inspirée des travaux sur la fluence et la disfluence orale, qui considèrent les hésitations dans la parole spontanée comme des moments où des

décisions et des choix sont faits par le locuteur. Plusieurs facteurs sont susceptibles d'influencer les pauses à l'écrit : des facteurs extralinguistiques (par exemple, l'humeur de l'individu), discursifs (le genre textuel), sémantiques (la familiarité sémantique), syntaxiques (la structure des phrases), lexicaux (la longueur des mots) et physiologiques (la production de la trace écrite). Des pauses longues seraient un indice de traitements longs et coûteux et des pauses brèves reflèteraient des traitements rapides et peu coûteux. Deux caractéristiques principales des pauses ont été étudiées : leur durée et leur localisation dans le texte.

Les pauses d'écriture assurent au moins deux fonctions (Schilperoord, 2002). Tout d'abord, elles permettent de préparer la suite du texte à produire. Les rédacteurs peuvent opérer à un niveau conceptuel pour planifier le contenu de leur texte ou à un niveau linguistique pour lexicaliser et linéariser les contenus. Les pauses à l'écrit permettent aussi de réviser le texte. À l'écrit le rédacteur n'est pas soumis à une pression temporelle : la situation de communication est asynchrone et le rédacteur dispose d'un temps plus long pour rédiger son texte et le transmettre à son ou ses lecteur(s). Il peut donc effectuer des pauses à loisir pour ajuster son texte aux objectifs communicationnels. D'autre part, la production écrite laisse une trace pérenne. De ce fait, alors qu'à l'oral les révisions qui font suite à des pauses portent généralement sur les parties de discours qui viennent d'être immédiatement émises, la révision à l'écrit peut porter sur des parties de texte produites bien plus antérieurement. Les rares travaux qui ont été conduits sur les relations entre la durée des pauses et la révision indiquent que les pauses qui précèdent une révision occupent environ 40 % de la durée totale de pause. Elles sont d'autant plus longues que les segments révisés sont longs et de haut niveau (van Waes & Schellens, 2003), mais le nombre et la durée des pauses semblent varier également selon la fonction que la révision assure : par exemple, les révisions sémantiques sont précédées de pauses plus longues que les révisions orthographiques (Matsuhashi, 1987). Les études portant sur les pauses d'écriture ont confirmé que des traitements de conceptualisation, de formulation linguistique mais aussi de révision peuvent être mis en œuvre de façon équiprobable durant les pauses de révision (Olive et al., 2009).

L'interprétation des pauses reste délicate compte tenu des nombreux facteurs qui peuvent affecter leurs durées et leurs fréquences. Les pauses étudiées dans la littérature sont celles pendant lesquelles des traitements cognitifs liés à la production sont mis en place. Ces pauses sont appelées pauses cognitives par comparaison aux pauses dites mécaniques qui sont la conséquence de contraintes mécaniques dues à l'outil ou au style d'écriture du rédacteur. En écriture manuelle, opérationnalisées comme des levers de crayon, certaines pauses relèvent en effet directement de la réalisation du tracé (espaces entre les mots, points sur les i, accents, etc.) et ne sont donc pas des moments de traitements cognitifs textuels. De même, l'écriture dactylographique conduit à des pauses systématiques entre chaque frappe.

Une première solution, qui est la plus communément adoptée mais qui pose plusieurs difficultés, consiste à ne pas prendre en compte ces pauses mécaniques qui sont généralement brèves et qui peuvent représenter 75% de la totalité des pauses (Cislaru & Olive, 2018) voire près de 90 % (Bouriga, 2020). Pour cela, un seuil est déterminé au-delà duquel on considère que les pauses sont sous la dépendance des opérations liées à la production du texte et non sous celles de la transcription graphique. Différentes valeurs ont été choisies en écriture manuscrite : 130 ms (Alamargot et al., 2007), 200 ms (Passerault, 1991), 250 ms (Olive & Kellogg, 2002, Olive et al., 2009), 2 secondes (Chenoweth & Hayes, 2001, 2003 ; Cislaru & Olive, 2018), ou en frappe au clavier : 1 seconde (Alves, Castro & Olive, 2009), 2 secondes (Alves et al., 2007), 5 secondes (Jansen et al., 1996). Décider d'un seuil de pause est donc délicat en raison de l'absence de justification théorique ou empirique.

D'autres approches ont essayé de catégoriser les pauses en fonction de leur durée. Une première se base sur la distribution des pauses et postule que cette distribution est composée de plusieurs sous-distributions, chacune renvoyant à un niveau de traitement spécifique. Trois

distributions ont été identifiées par Baijjen et al. (2012) : une première relative au processus de récupération des mots (65% des pauses), une deuxième renvoyant aux processus à la frontière des phrases (26%) et une troisième représentant le processus de planification (2% des pauses). Selon ces auteurs, confronter la question de recherche à ces trois distributions peut orienter la définition du seuil, voire l'utilisation de seuils multiples dans une même étude pour mettre en évidence des effets des différents traitements. Chenu et al. (2014) décrivent différentes approches pour traiter cette question du seuil de pause. Ils proposent par exemple de repérer des points de rupture, des disfluences, dans la dynamique d'écriture. Des recherches spécifiquement dédiées à cette question doivent être poursuivies.

Une autre difficulté posée par la détermination d'un seuil de pause est la prise en compte des caractéristiques individuelles des rédacteurs. Fixer un seuil de pause identique pour tous revient à faire l'hypothèse implicite qu'une contrainte, quelle qu'elle soit, entraîne des pauses mécaniques de même durée, quel que soit le rédacteur, et en particulier quel que soit son style d'écriture (cursive, script) ou son habileté d'écriture ou de frappe au clavier. Alors qu'un rédacteur avec une écriture cursive ne fera presque pas de pauses intra-mot, un rédacteur avec une écriture en script ou avec un clavier fera une pause entre chaque lettre et chaque frappe. Définir un seuil individuel en fonction du rédacteur et de la situation de rédaction semble donc indispensable. Dans sa thèse, qui comparait l'écriture manuscrite et la frappe au clavier, Bouriga (2020) a été confrontée au problème de l'individualisation des seuils de pause, mais aussi à celui de la comparabilité des phénomènes observés selon les situations d'écriture. Elle comparait en effet des situations de production manuscrite et dactylographique chez un même rédacteur. Pour pallier ces difficultés, Bouriga (2020) a développé une méthode prenant en compte les différences intra- et inter-individuelles tout en autorisant un certain degré de comparabilité avec les études qui ont utilisé un seuil de 2 secondes. La solution adoptée pour calculer ce seuil individuel a été de retenir comme valeur de référence le seuil le plus fréquemment utilisé dans la littérature, soit 2 secondes. Le percentile lui correspondant dans la distribution de l'ensemble des pauses de tous les participants de chaque étude de sa thèse a été identifié. Puis ce percentile a été appliqué à la distribution des pauses de chaque participant, dans chaque modalité de rédaction, pour repérer la durée de pause qui était associée et qui devait servir de seuil individuel. Cette méthode présente un double avantage : elle prend en compte les différences interindividuelles et intra-individuelle en termes d'habiletés d'écriture et de frappe et elle permet de comparer les résultats avec ceux des études antérieures. Il faut noter que les patrons de pause obtenus avec cette méthode se sont révélés identiques à ceux observés dans les études précédentes ayant utilisé un seuil unique.

3.2. Un bilan rapide des travaux

Les premiers travaux sur les pauses ont tenté de mettre en évidence une unité de planification du langage en faisant l'hypothèse que les pauses sont des moments de réflexion visant à préparer le segment de discours suivant. En effet, si les pauses sont des moments privilégiés de planification du discours, alors elles devraient être plus fréquemment localisées aux frontières des "unités" de planification du langage. Toutefois, la nature des unités invoquées diffère selon l'approche des chercheurs. Pour certains, les pauses interviendraient principalement à des frontières sémantiques. Butterworth (1975) montre ainsi que le discours oral est composé de cycles alternant des phases fluides et d'hésitations, organisées autour de configurations sémantiques. Lors des phases fluides de production, c'est-à-dire entrecoupées par peu de pauses, les locuteurs planifieraient leur discours essentiellement au niveau morphosyntaxique, alors que lors des phases disfluides ils planifieraient davantage les aspects conceptuels de leur discours. De plus, les phases fluides portant souvent sur la production de plusieurs phrases, la planification conceptuelle englobe alors plusieurs énoncés.

D'autres auteurs ont mis en avant la fonction de planification syntaxique des pauses. Il est en effet apparu dans les travaux mettant en parallèle les pauses et leur localisation dans la structure syntaxique du discours que celles-ci étaient plus fréquentes aux frontières des unités grammaticales et plus longues en fonction de la structure qu'elles précèdent. Van Hell et al. (2008) ont examiné la durée des pauses avant divers types de propositions plus ou moins complexes. Ils montrent que les pauses sont plus longues à des niveaux hiérarchiques élevés et qu'à un même niveau hiérarchique, la durée des pauses varie en fonction de la nature des propositions qui les suivent. Chez les adultes comme chez les enfants, les pauses les plus longues précédaient les propositions principales. La durée des pauses d'écriture augmenterait donc avec la complexité syntaxique des unités langagières qui les suivent. Foulin (1998) a aussi tenté d'expliquer la durée des pauses en fonction de leur localisation syntaxique. Il montre que la fréquence des pauses peut témoigner de la planification macro- et microstructurale du discours. La durée des pauses entre paragraphes relèverait ainsi principalement de la gestion des connaissances. La durée des pauses entre phrases résulterait des processus syntaxiques et lexicaux qui pourraient intervenir conjointement à la planification conceptuelle. Les pauses inter-propositionnelles témoigneraient principalement de la planification linguistique du discours, c'est-à-dire de l'organisation syntaxique du discours. Enfin, les pauses intra-propositionnelles seraient quant à elles affectées par la prévisibilité des items lexicaux.

D'autres études ont montré que les pauses sont également sensibles aux différentes contraintes rédactionnelles (i.e., connaissance du thème ou du type de texte, destinataire, outils d'écriture, coût des processus rédactionnels, structure syntaxique suivant la pause, etc.). Par exemple, la durée de pause est plus grande lorsque le texte à rédiger est argumentatif plutôt que narratif. L'augmentation de la durée des pauses serait liée à la plus grande complexité de planification du texte argumentative, alors que le texte narratif est considéré comme maîtrisé avec un schéma de rédaction bien internalisé. Cette augmentation du nombre et de la durée des pauses a également été observée avec la rédaction sur ordinateur par rapport à l'écriture manuscrite (van Waes & Schellens, 2003). Par ailleurs, une faible maîtrise de la transcription, par exemple en demandant une rédaction de texte en lettres majuscules, ou de la formulation linguistique entraîne une augmentation du nombre de pauses.

Pour mieux comprendre la gestion en temps réel de la production écrite, il semble indispensable d'étudier également les variations de débit. Par exemple, Foulin (1998) a observé des résultats semblables dans une tâche de production de phrases à deux propositions. Alors que la durée des pauses intra-propositionnelles chutait entre la première et la deuxième proposition chez des adultes, elle était plus élevée dans la seconde proposition chez des enfants. De plus, les rédacteurs adultes écrivaient d'autant plus vite qu'ils se rapprochaient de la fin de la phrase alors que les enfants ne modulaient pas leur vitesse d'exécution. Foulin conclut que, chez les adultes, la planification conceptuelle et linguistique se déroule au cours de la transcription de la première proposition. Toutefois, cette planification serait partielle, elle se poursuivrait au cours de la transcription des propositions suivantes (voir aussi Chanquoy et al., 1990).

En somme, les questions soulevées par l'analyse des pauses et des débits témoignent de la complexité du système de production du langage. Outre les facteurs extra-cognitifs, les durées et fréquences de pauses sont déterminées par la pluralité des processus cognitifs qui interviennent lors de la production verbale. En général, les recherches font apparaître un impact des frontières syntaxiques sur les pauses. Celles-ci sont d'autant plus longues que l'unité à produire est élevée (paragraphe, phrase, proposition, syntagme) dans l'organisation syntaxique du discours et que le discours à produire est complexe. Les débuts de phrases et de propositions seraient des moments privilégiés de planification conceptuelle et linguistique. L'analyse des pauses et des débits révèle également un recouvrement entre ces opérations et celles d'exécution (articulation ou transcription). De plus, la durée des pauses est largement influencée par les

activités de révision ce qui souligne encore davantage le caractère multi-déterminé des pauses et fait ressortir la nécessité d'interpréter les pauses en relation systématique avec les modèles théoriques de la production du langage.

4. À l'interface processus-produits : les périodes de transcription et les jets textuels

Les périodes d'exécution (en anglais *bursts of language*) sont des moments de transcription graphique ininterrompus, c'est-à-dire délimités par deux pauses. Il faut noter que la définition d'une période d'exécution peut varier d'une recherche à l'autre. Par exemple, dans les travaux de Chenoweth et Hayes (e.g., 2001, 2003, Hayes et Chenoweth, 2006) ou d'Alves et collaborateurs (e.g., 2016, 2017), les périodes d'exécution doivent contenir au moins un mot. Pour Cislaru et Olive (2018), une période d'exécution n'est définie opérationnellement que par ses bornes. Ils observent d'ailleurs que nombre de périodes d'exécution qui contiennent moins qu'un mot informent sur le processus de textualisation. Par exemple de nombreuses périodes d'exécution ne contiennent, en français, qu'une lettre assurant un marquage morphologique (par exemple un « s » pluriel).

4.1. Les périodes de transcription

Les périodes d'exécution ont été longtemps ignorées parce qu'elles étaient considérées comme des moments de simple transcription graphique. Or, les rédacteurs habiles peuvent transcrire un segment de texte et simultanément préparer le segment suivant ou réviser le ou les segments précédents (voir plus haut). Plusieurs recherches ont ainsi montré l'impact de variations liées aux caractéristiques individuelles des rédacteurs ou aux contextes et situations d'écriture sur les périodes d'exécution.

En premier lieu, ces travaux confirment l'impact de la maîtrise de la transcription sur les périodes d'exécution : elles sont plus courtes lorsque la transcription est peu maîtrisée. C'est le cas en écriture manuscrite chez les enfants et chez les adultes comme en écriture dactylographique. On peut d'ailleurs noter que des entraînements à la transcription chez des enfants aboutissent à la production de périodes d'exécution plus longues (Alves et al. 2016 ; Limpo & Alves, 2017). En second lieu, l'impact des habiletés de formulation linguistique a aussi été démontré. Ainsi, les rédacteurs experts rédigent avec des périodes d'exécution plus longues que celles des rédacteurs novices (en anglais, voir Kaufer et al., 1986 ; en mandarin, voir Friedlander, 1989). De même, lorsque les étudiants rédigent un texte dans leur langue maternelle, les périodes d'exécution comportent plus de mots que lorsque ces mêmes étudiants composent en langue seconde (Chenoweth & Hayes, 2001). Similairement, les périodes d'exécution sont plus longues chez des rédacteurs plus avancés dans l'apprentissage d'une langue seconde que chez les rédacteurs moins avancés (voir aussi Chenoweth & Hayes, 2003). Les enfants qui présentent un trouble spécifique du langage produisent aussi avec des périodes de transcription plus courtes que les enfants au développement langagier typique (Connelly et al., 2012). Enfin Bouriga (2020) n'a pas montré d'impact de la complexité des traitements de conceptualisation ou de révision sur les périodes d'exécution. Les paramètres temporels des périodes d'exécution semblent donc dépendre essentiellement du niveau d'automatisation et de maîtrise de la transcription graphique et de la formulation linguistique.

La compréhension des périodes d'exécution a été précisée en les catégorisant selon leur fonction textuelle. Deux types ont été plus étudiés : les périodes dites de production et celles dites de révision. Les premières incrémentent le texte sur son bord droit, les secondes interviennent sur le texte déjà produit. Dans les termes de Cislaru et Olive (2018), lorsqu'une période d'exécution suit la linéarité chronologique et topologique, elle est catégorisée comme une production. En revanche, lorsqu'une période d'exécution respecte la linéarité chronologique mais pas celle topologique, elle est catégorisée comme une révision. Les

périodes d'exécution de révision sont généralement moins nombreuses et plus courtes que celles de production. Leur vitesse de frappe est également plus rapide et elles sont précédées de pauses plus longues.

Baijjen et al. (2012) ont identifié 13 catégories de périodes d'exécution selon que les périodes d'exécution étaient de production ou de révision (immédiates, distantes, ou mixtes), selon leur localisation (e.g., point d'insertion), et le type de modification qu'elles créent (e.g., insertion, remplacement). À ce jour, leur travail n'a pas été poursuivi et il semble que la distinction la plus heuristique soit celle de production/révision, en distinguant deux types de périodes de révision : celles qui portent sur le segment textuel produit lors de la dernière période de production, il s'agit alors d'une révision dite immédiate, et celles qui exigent un retour sur le texte au-delà de la dernière période d'exécution, on parle alors de période de révision distante.

4.2. Des périodes de transcription aux jets textuels

Les pauses, utilisées pour identifier les périodes d'exécution, sont aussi des marqueurs de segmentation textuelle : chaque période d'exécution isole un segment textuel, un jet textuel. L'identification des jets textuels s'appuie donc sur une segmentation opérée à partir du processus de rédaction, c'est-à-dire pendant que le texte est construit. On peut parler de segmentation spontanée dans la mesure où les pauses sont le fait des rédacteurs. Bien entendu, le choix d'un seuil de pause dépend d'une décision du chercheur et le contenu et la fonction qui peuvent être attribués à un jet textuel varient fortement selon la valeur du seuil sélectionné.

Cislaru et Olive (2018) proposent de décrire les jets textuels en termes de structure interne, de relations de dépendance, de formes et de fonctions ainsi que de mécanismes de constitution (par exemple via des processus de *chunking*). Ils observent que les jets textuels montrent une grande hétérogénéité en termes de longueur et de composants langagiers : une très grande majorité de jets textuels est constituée de séquences incomplètes ou non finies syntaxiquement, dont la borne droite engage un schéma syntagmatique qui sera complété par le jet textuel suivant. Ces auteurs tentent aussi de décrire les régularités formelles, sémantiques et relationnelles des jets textuels, chaque régularité pouvant signaler un statut linguistique. Les régularités identifiées montrent que la production écrite est sous-tendue par des logiques constructionnelles, dans des formats préconstruits non saturés. Par conséquent, certains jets textuels pourraient indiquer un degré d'automatisme, qui se traduirait non par une génération mais par une récupération en mémoire à long terme d'au moins une partie des informations (Cislaru & Olive, 2017). Cislaru et Olive (2018) proposent également de décrire ces jets textuels en tant qu'unités agglomérant des éléments autour d'un attracteur de jonction, lequel pourrait sémiotiser une relation ou une accréation sémantique au niveau textuel (connecteur, anaphore, coordination).

L'analyse des jets textuels permet ainsi de lever le voile sur ce qui est invisible à l'observation : le processus de textualisation. On peut observer les formes linguistiques s'agencer et s'associer les unes aux autres d'une manière donnée, à des rythmes plus ou moins réguliers, en laissant entrevoir des schémas, des constructions, des séries de récurrences. Cette dynamique de la performance offre un angle de vue différent sur les unités linguistiques et les constructions qui les accueillent, à tous les paliers analytiques : syntagmes, propositions, phrases, séquences, textes. Enfin, la segmentation comportementale des jets textuels impose un regard interdisciplinaire.

5. Conclusion

Ce chapitre visait à présenter les bases théoriques et opérationnelles de l'étude en temps réel de la production écrite. Centré sur les méthodes et variables dépendantes spécifiques à l'écrit, comme les pauses ou les débits de production, l'étude en temps réel permet de décrire les

traitements impliqués lors de la production (et corollairement la nature des représentations engagées) ainsi que la façon dont ces traitements sont coordonnés. Ces mesures et leur interprétation reposent sur les postulats de la chronométrie mentale : il est possible de mesurer la durée des processus cognitifs, et leur durée reflèterait la complexité des traitements.

Ces méthodes ont permis de montrer qu'une partie des traitements de production écrite peuvent être mis en œuvre simultanément, et en particulier les processus centraux et de transcription graphique chez les rédacteurs experts. Un autre trait de la gestion en temps réel est la possibilité qu'ont les rédacteurs habiles d'adapter leur mode de fonctionnement aux demandes de la production : ils peuvent ainsi passer d'un mode parallèle lorsque les coûts le permettent à un mode plus séquentiel lorsque ces coûts augmentent.

En rendant possible l'analyse de la temporalité de la production écrite, l'écriture enregistrée donne également accès à la dynamique du texte « en train de se faire ». Cislaru et Olive (2018) ont ainsi proposé une théorie de la textualisation basée sur la description de la structure interne des jets textuels et sur la façon dont ces derniers sont articulés pour construire un texte.

Bien sûr, d'autres techniques en temps réel peuvent être utilisées. On peut penser par exemple aux verbalisations concomitantes, à l'analyse des temps de réaction secondaires ou encore aux techniques d'imagerie cérébrale dont certaines permettent de suivre précisément le décours temporel des opérations cognitives de production du langage écrit (voir par exemple les travaux de Perret et collaborateurs sur la cartographie de l'activité électro-encéphalographique des rédacteurs, Perret et al., 2014). Ces techniques n'ont pas été présentées car elles ne sont pas spécifiques à la production écrite. Nul doute qu'elles apporteront des éclairages nouveaux sur la production écrite.

Points à retenir

- L'approche en temps réel de la production écrite analyse le fonctionnement cognitif des rédacteurs et la dynamique de la production écrite à partir d'enregistrements de l'activité d'écriture.
- Des données comme la durée des pauses ou les vitesses d'écritures sont analysées pour mettre à jour les stratégies cognitives des rédacteurs et le processus de textualisation
- La durée des pauses dépend généralement de leur localisation : les unités les plus hautes et les plus complexes aboutissent à des pauses plus longues.
- Les périodes d'exécution renvoient aux phases de transcription interrompues ; elles sont principalement affectées par les traitements de formulation et de transcription.
- Les segments de texte identifiés lors des périodes de production renseignent sur le processus de textualisation ; la plupart de ces segments sont syntaxiquement non saturés.

Pour aller plus loin

- Sur la production de mots :

Perret, C., & Olive, T. (2019). *Spelling and Writing Words : Theoretical and Methodological Advances*. Brill. <https://doi.org/10.1163/9789004394988>

- Sur la production de textes :

La collection d'ouvrages pluridisciplinaire, *Studies in Writing*, éditée par Raquel Fidalgo et Thierry Olive et publiée par Brill : <https://brill.com/view/serial/SIW>

- Sur l'enregistrement de l'écriture :

Lindgren, E., & Sullivan, K. (Eds.). (2019). *Observing writing: Insights from keystroke logging and handwriting*. In R. Fidalgo & T. Olive (Series Eds), *Studies in Writing* (Vol. 38). Brill. <https://doi.org/10.1163/9789004392526>

Bibliographie

- Alves, R. A., & Limpo, T. (2015). Progress in Written Language Bursts, Pauses, Transcription, and Written Composition Across Schooling. *Scientific Studies of Reading*, 19(5), 374-391. <https://doi.org/10.1080/10888438.2015.1059838>
- Alves, R. A., Castro, S. L., de Sousa, L., Strömqvist, S. (2007). Influence of typing skill on pause-execution cycles in written composition ». In M. Torrance, L. van Waes, & D. Galbraith (Eds), *Writing and Cognition : Research and Applications* (pp. 55–65). Elsevier Publishers.
- Alves, R. A., Leal, J. P., & Limpo, T. (2019). HandSpy: A new tool to study handwriting in real time. In E. Lindgren and K. P. H. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 50-70). Brill.
- Alves, R. A., Limpo, T, Fidalgo, R., Carvalhais, L., Pereira, L. Á., & Castro, S. L. (2016). The impact of promoting transcription on early text production : Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology*, 108(5), 665–679
- Baaijen, V. M., Galbraith, D., & Glopper, K. de. (2012). Keystroke Analysis : Reflections on Procedures and Measures. *Written Communication*, 29(3), 246-277. <https://doi.org/10.1177/0741088312451108>
- Berg, T. (2002). Slips of the typewriter key. *Applied Psycholinguistics*, 23, 185-207.
- Bonin, P., & Fayol, M. (2002). The written production of isolated words. In T. Olive & C. M. Levy (Eds), *Contemporary tools and techniques for studying writing* (pp. 119-130). Kluwer Academic Publishers.
- Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: different tasks, different pathways? *The Quarterly Journal of Experimental Psychology*, 68(7), 1268–1294. <https://doi.org/10.1080/17470218.2014.978877>.
- Bouriga, S. (2020). *Papier-crayon vs. écran-clavier : Effets sur le coût cognitif et sur la dynamique de la production de textes* [Thèse de doctorat de l'université de Poitiers]. Thèse.fr. <https://www.theses.fr/2020POIT5003>
- Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4(1), 75–87.
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, 37, 243-297.
- Cattell, J.M. (1886). The Time Taken up by Cerebral Operations, *Mind*, 11, 220-242.
- Chanquoy, L., Foulon, J.-N., & Fayol, M. (1990). The on-line management of short text writing by children and adults, *CPC/European Bulletin of Cognitive Psychology*, 10, 513-540.
- Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18, 80–98. <https://doi.org/10.1177/>
- Chenoweth, N. A., & Hayes, J. R. (2003). The Inner Voice in Writing. *Written Communication*, 20(1), 99-118. <https://doi.org/10.1177/0741088303253572>
- Chenu, F., Pellegrino, F., Jisa, H., & Fayol, M. (2014). Interword and intraword pause threshold in writing. *Frontiers in Psychology*, 5, 189. <https://doi.org/10.3389/fpsyg.2014.00182>
- Chesnet, D., Solier, C., Bordas, B., & Perret, C (2022). A quick briefing on the new version of Eye and Pen (version 3.01): news and update. *Reading and Writing*. <https://doi.org/10.1007/s11145-022-10267-5>.

- Cislaru, G., & Olive, T. (2017). Quel niveau de pré-construction ? Segments répétés, jets textuels et autres routines. *Corpus*, 17, 61-89.
- Cislaru, G. & Olive T. (2018). *Le processus de textualisation*. De Boeck.
- Connelly, V., Dockrell, J., Walter, K., & Critten, S. (2012). Predicting the quality of composition and written language bursts from oral language, spelling, and handwriting skills in children with and without Specific Language Impairment. *Written Communication*, 29(3), 278302. <https://doi.org/10.1177/0741088312451109>
- Crump, M. J. C., & Logan, G. D. (2010). Hierarchical control and skilled typing: Evidence for word-level control over the execution of individual keystrokes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(6), 1369–1380. <https://doi.org/10.1037/a0020696>
- Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Language*, 32, 1330-1140.
- De Sousa, O., Olive, T. & Lambert E. (2019). Writing before speaking modifies speech production: Evidence from a non-color-word Stroop task. *Experimental Psychology*. 66, 126-133. <https://doi.org/10.1027/1618-3169/a000434>
- Ellis, A. W. (1988). Normal writing processes and peripheral acquired dysgraphias. *Language and Cognitive Processes*, 3, 99-127.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116124. <https://doi.org/10.3758/BF03195503>
- Foulin, J.-N. (1995). Pauses et débits : les indicateurs temporels de la production écrite, *L'Année Psychologique*, 95, 483-504.
- Foulin, J.-N. (1998). To what extent does pause location predicts pause duration in adult's and children's writing ? *Current Psychology of Cognition*, 17, 601-620
- Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication*, 23(2), 135–149. <https://doi.org/10.1177/0741088306286283>
- Jansen, D., Van Waes, L., & Van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In C.M. Levy & S. Ransdell (Eds.), *The Science of writing: Theories, methods, individual differences and applications* (pp. 233–250). Lawrence Erlbaum Associates.
- Kandel, S., & Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition? *Cognition*, 136, 325-336.
- Kandel, S., Alvarez, C. J., & Vallée, N. (2006). Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 19-31.
- Kaufert, D. S., Hayes, J. R. & Flower, L. (1986). Composing written sentences, *Research in the Teaching of English*, 20(2), 121140.
- Lambert, E., Alamargot, D., Larocque, D., & Caporossi, G. (2011). Dynamics of the spelling process during a copy task: effects of regularity and frequency. *Canadian journal of experimental psychology*, 65(3), 141–150. <https://doi.org/10.1037/a0022538>
- Leijten, M. & Van Waes, L. (2006). Inputlog: New Perspectives on the logging of on-Line writing ». In G R.ijlaarsdam (Series éd.) *Studies in Writing*, Vol. 18, and Kirk P. H. Sullivan, & E. Lindgren (Vol. Eds.), *Computer Keystroke Logging and Writing: Methods and Applications* (pp.73-94). Elsevier.
- Limpo, T., & Alves, R. A. (2017). Written Language Bursts Mediate the Relationship Between Transcription Skills and Writing Performance. *Written Communication*, 34(3), 306–332. <https://doi.org/10.1177/0741088317714234>

- Lindgren, E., & Sullivan, K. (Eds.). (2019). *Observing writing: Insights from keystroke logging and handwriting*. In R. Fidalgo & T. Olive (Series Eds.), *Studies in Writing* (Vol. 38). Brill. <https://doi.org/10.1163/9789004392526>
- Logan, G. D. (2018). Automatic control: How experts act without thinking. *Psychological Review*, 125(4), 453–485. <http://doi.org/10.1037/rev0000100>
- Maggio, S., Lété, B., Chenu, F. et al (2012). Tracking the mind during writing: immediacy, delayed, and anticipatory effects on pauses and writing rate. *Reading and Writing*, 25, 2131–2151. <https://doi.org/10.1007/s11145-011-9348-1>
- Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, 6, 173-194.
- Olive, T. (2022). How do novice and skilled writing engage working memory? In J. W. Schweiter & J. Cheng (Eds.), *Cambridge Handbook of Working Memory and Language* (pp. 504-526). Cambridge University Press. <https://doi.org/10.1017/9781108955638.028>
- Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology*, 21, 758–785. <https://doi:10.1080/09541440802079850>
- Olive, T. & Cislaru, G. (2015). Linguistic forms at the process-product interface: Analyzing the linguistic content of bursts of production. In G. Cislaru (Ed.), *Writing(s) at the crossroads: the process/product interface* (pp. 99-123). John Benjamins.
- Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory and Cognition*, 30, 594-600. <https://doi.org/10.3758/BF03194960>
- Passerault, J.M. (1991). L'analyse en temps réel de l'activité de résumé : une étude des temps de pauses au cours de la "re-mise" en texte. In M. Charolles & A. Petitjean (Eds.), *Le résumé de texte : aspects linguistiques, sémiotiques, psycholinguistiques et automatiques* (pp. 207-219). Klincksieck
- Perret, C., & Laganaro, M. (2011). Comparison of Electrophysiological Correlates of Writing and Speaking: A Topographic ERP Analysis. *Brain Topography*, 25(1), 6472. <https://doi.org/10.1007/s10548-011-0200-3>
- Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written picture naming using a topographic ERP analysis. *Brain and Language*, 135, 20-31. <https://doi.org/10.1016/j.bandl.2014.04.006>
- Perret, C., & Olive, T. (2019). *Spelling and Writing Words: Theoretical and Methodological Advances*. Brill. <https://doi.org/10.1163/9789004394988>
- Sausset, S., Lambert, E., Olive, T., & Larocque, D. (2012). Processing of syllables during handwriting: Effects of graphomotor constraints. *Quarterly Journal of Experimental Psychology*, 65(10), 18721879. <https://doi.org/10.1080/17470218.2012.715654>
- Schilperoord, J. (2002). On the Cognitive Status of Pauses in Discourse Production. In C. M. Levy & T. Olive (Eds.), *Contemporary Tools and Techniques for Studying Writing* (pp. 6087). Kluwer Academic Publisher. https://doi.org/10.1007/978-94-010-0468-8_4
- Séraphin-Thibon, L. S., Gerber, S., & Kandel, S. (2018). Analyzing Variability in Upper-Case Letter Production in Adults. In C. Perret & T. Olive (Eds.) *Spelling and writing words: Theoretical and Methodological Advances* (pp. 163-178). Brill. <https://doi.org/10.1016/j.actpsy.2017.12.001>
- Strömqvist, S., & Malmsten, L. (1998). *Scriptlog Pro 1.04 – User's Manual*. Technical report. Göteborg University, Department of Linguistics.
- Yamaguchi, M., & Logan, G. D. (2014). Pushing typists back on the learning curve: Contributions of multiple linguistic units in the acquisition of typing skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1713–1732. <https://doi.org/10.1037/xlm0000026>

- van Galen, G.P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science*, 10, 165-191.
- van Hell, J. Verhoeven, L., van Beijsterveldt, L. (2008). Pause time patterns in writing narrative and expository texts by children and adults. *Discourse Processes* 45, 406–427.
- van Waes, L, & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics*, 35(6), 829853. [https://doi.org/10.1016/s0378-2166\(02\)00121-2](https://doi.org/10.1016/s0378-2166(02)00121-2)
- Wengelin, A., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods*, 41(2), 337351. <https://doi.org/10.3758/brm.41.2.337>