



Orientales

Langues, cultures, pratiques

Sous la direction de

Madeleine VOGA

& Mariarosaria GIANNINOTO

DM

東方
الشرق

3

2025

Orientales
n° 3
Langues, cultures, pratiques

Numéro coordonné par
Madeleine VOGA & Mariarosaria GIANNINOTO

Revue dirigée par
Mariarosaria GIANNINOTO & Abdenbi LACHKAR

Décembre 2025
PRESSES UNIVERSITAIRES DE LA MÉDITERRANÉE

Revue Orientales

Fondée à l'initiative de chercheurs sinisants et arabisants de l'université de Montpellier Paul-Valéry, cette revue se propose de devenir un lieu d'échange sur les études orientales. Ces dernières années ont vu l'essor économique des pays de l'aire extrême-orientale et de l'Asie méridionale et de nombreux changements sociaux et économiques ont marqué les pays du monde arabe ou encore d'Asie centrale. Cela s'accompagne d'un développement considérable de l'enseignement des langues orientales, ainsi que du développement correspondant des études sur les langues et les cultures de ces aires bénéficiant d'une institutionnalisation croissante. Orientales s'inscrit dans ce contexte et se propose de contribuer au développement des recherches sur les langues, les cultures et les sociétés orientales. La revue pourra accueillir non seulement des études centrées sur ces aires géographiques et culturelles et sur les interactions de ces aires entre elles, mais aussi sur les interactions, les rencontres et les transferts entre ces aires et les pays européens.

Les articles sont publiés en français ou en anglais. Les articles sont soumis à une évaluation anonyme par deux membres du comité scientifique ou par d'autres spécialistes.

Directeurs de publication

Mariarosaria GIANNINOTO (UPMV) et Abdenbi LACHKAR (UPMV)

Comité de rédaction

Nancy BALARD (UPMV), Solange CRUEILLÉ (UPMV), Mariarosaria GIANNINOTO (UPMV), Rima LABBAN (UPMV), Abdenbi LACHKAR (UPMV), Félix Jun MA (UPMV).

Comité scientifique

Brahim CHAKRANI (Michigan State University, USA), Laurence DENOZ (Université de Lorraine), Samia FERHAT (Université Paris Nanterre), Shkooch HOSEINI (IHCS-Téheran), Tomoko HIGASHI (Université Grenoble Alpes), Hayssam KOTOB (Université Libanaise), Zahra MOBALLEGH (IHCS-Téheran), Bruno PAOLI (Université Lyon 2), Barbara RAHMA (Université de Fès), Dris SOULAIMANI (San Diego State University), Madeleine VOGA (UPV), Frédéric WANG (INALCO, Paris), Cécilia YU XINYUE (INALCO, Paris), Ying ZHANG-COLIN (Université Paris Nanterre).

L'expertise ponctuelle de collègues extérieurs au comité scientifique est régulièrement sollicitée pour des articles traitant de langues et de domaines qui ne rentrent pas dans les champs d'expertise des membres du comité.

Les propositions d'articles sont à envoyer aux adresses suivantes : mariarosaria.gianninoto@univ-montp3.fr, abdenbi.lachkar@univ-montp3.fr

Site de la revue en ligne

www.pulm.fr

Illustration de couverture : Photomontage PULM, 2025.

ISSN 3003-2105 — ISBN 978-2-36781-589-3

Tous droits réservés, PULM, 2025.

Sommaire/Contents

Madeleine VOGA & Mariarosaria GIANNINOTO	
<i>Introduction. Lexique et grammaire en usage :</i>	
<i>données et représentations</i>	5
Xiaoliang LUO	
<i>Enseignement de la prononciation du chinois aux étudiants</i>	
<i>francophones : une méthode basée sur la phonologie</i>	15
Hélène GIRAUDO, Serena DAL MASO & Sabrina PICCININ	
<i>Conscience morphologique et lecture chez les élèves en Réseau</i>	
<i>d'Éducation Prioritaire : une étude pilote en fin de CE1</i>	35
Nikolaos NTAGKAS	
<i>Associating Form and Meaning in Morphological Theory:</i>	
<i>Experimental Data From Modern Greek.</i>	61
A. FLIATOURAS, M. VOGA & A. ANASTASSIADIS-SYMÉONIDIS	
<i>L'étymologie populaire comme mécanisme de changement linguistique :</i>	
<i>apport des données du grec et implications pour le lexique mental</i>	79
Alberto BRAMATI	
<i>La traduction en italien du pronom ça sujet</i>	107
Guo JING	
<i>Enseigner la politesse chinoise aux apprenants francophones :</i>	
<i>représentations et difficultés d'appropriation.</i>	127
Madeleine VOGA	
<i>Lexique mental bi-plurilingue : les données chronométriques</i>	
<i>inter-alphabet au service d'une approche unificatrice</i>	141
Enhao LEGER-ZHENG	
<i>Les apprenants de chinois langue étrangère repèrent-ils les mots</i>	
<i>inconnus dans leur lecture ? Une étude d'oculométrie auprès</i>	
<i>d'apprenants francophones.</i>	169
<i>Auteurs</i>	195

Introduction

Lexique et grammaire en usage : données et représentations

Madeleine VOGA & Mariarosaria GIANNINOTO
ReSO, Université de Montpellier Paul Valéry

Ce numéro se situe au croisement de plusieurs domaines ou (sous-)disciplines de la linguistique appliquée (Stegu 2011 ; Linn, Candel & Léon 2011), incluant des articles sur l'acquisition, le traitement et l'enseignement/apprentissage des langues. Il recueille une partie des contributions issues du colloque « Plurilexgram — Lexique et grammaire en usage : données et représentations », qui s'est tenu à Montpellier les 9 et 10 novembre 2023, ainsi que d'autres contributions spontanées portant sur la compétence lexicale, le changement linguistique, la description des langues et leur enseignement/apprentissage, avec un accent mis sur l'étude de langues et/ou de combinaisons linguistiques habituellement peu étudiées.

La compétence lexicale du locuteur et la connaissance des mots de sa ou ses langues, auxquelles on se réfère par le terme *lexique mental*, ainsi que les rapports que ce dernier entretient avec la grammaire ont connu, ces dernières décennies, des bouleversements. Pour ne citer que cet exemple, la distinction (Kiparsky, 1982) entre « lexique permanent », où l'on trouve les lemmes/unités lexicales sans structure interne, et le « lexique dynamique », où l'on trouve les processus morphologiques, *i. e.*, la dérivation, composition et flexion des mots, a cédé la place à une approche plus unifiée, où les processus ne sont pas isolés du support sur lequel ils opèrent et où la grammaire (et/ou la morphologie) fonctionne parallèlement à la syntaxe (*e. g.* Booij, 2002).

Si le progrès et les avancées dans le domaine sont indéniables, il convient de souligner une certaine dissymétrie entre l'immense masse de données accumulées pour certaines langues *versus* la quantité limitée d'études et de données consacrées à d'autres. Des *habitus* en matière de recherche ont également contribué à la sur-représentation de certaines langues, dont l'anglais, qui finissent par imposer une sorte de norme implicite, en matière de données quantitatives mais égale-

ment qualitatives. Compte tenu de l'importance accordée aux différents types de formalisation, en particulier pour les cognitivistes, ainsi qu'aux processus mentaux (pour une tentative récente portant sur la morphologie, cf. Jackendoff & Audring 2020), une partie des études linguistiques a perdu de son champ de recherche la diversité des langues. Ceci a comme corollaire la sous-représentation de langues et/ou de combinaisons linguistiques moins fréquentes, voire la mise à l'écart de faits de langue, en raison de leur éloignement par rapport à cette norme implicite. Comme le fait remarquer Rastier :

Si personne parmi les cognitivistes ne conteste l'autonomie, peut être relative de la pensée, tous dénie celle du langage. C'est à cette dénégation que la sémantique doit d'occuper dans la linguistique un rôle central : en effet, le sens est identifié — comme il l'a toujours été jusqu'à Saussure — à des représentations. En outre, les représentations sont généralement réduites à des concepts.
(Rastier, 1993 : 162)

Ce numéro se donne deux objectifs.

Il se propose tout d'abord de contribuer au développement, récent et pour le moment relativement timide sur le plan international, des travaux en français portant sur des langues autres que celles traditionnellement étudiées, *i. e.*, ce numéro s'inscrit ainsi directement dans le développement de travaux sur les langues « autres ». Cette entreprise, à laquelle nous espérons contribuer, peut constituer une avancée, dans la mesure où elle permettra aux chercheurs de disposer d'une assise plus solide pour l'étude d'objets tels que l'usage, la structure et le fonctionnement du lexique mental, notamment à travers l'étude des correspondances et différences interlinguistiques. Cette perspective s'impose d'elle-même dès lors que l'on s'intéresse à la diversité des langues, en termes de faits (description), de processus ou encore de supports. Cette entreprise exige néanmoins de prendre certaines précautions, la première étant liée à la variabilité inhérente aux langues et à leur grande diversité. Outre celle-ci, la grande variabilité en termes de profils de locuteurs et de conditions d'acquisition, peut poser des défis particuliers lorsqu'il s'agit de décrire le lexique et/ou la grammaire d'une langue. Cela se reflète à la fois sur le plan formel ainsi que sur les plans sémantique et conceptuel. Pour ce qui concerne les « niveaux entrants » de représentation et de traitement, la graphie et la phonie influencent, voire définissent à la fois la perception et la production, aussi bien durant le traitement de la part du sujet natif, que durant l'acquisition et apprentissage chez le locuteur non-natif. Pour ce qui concerne le niveau sémantique et conceptuel, des défis se posent lorsqu'il s'agit de collecter, analyser et interpréter des données linguistiques qui se veulent objectives et rigoureuses. À ces niveaux de traitement et de représentation s'ajoute le niveau pragmatique, qui, à l'instar de la morphologie et du lexique, se trouve à intersection du formel et du sémantique-conceptuel.

Ce numéro se propose aussi de faire avancer la réflexion sur ce que l'on peut appeler diasystème. Weinreich (1954 : 390) emploie cette notion pour expliquer les correspondances régulières entre des structures différentes dans des variétés étroitement liées. Nous adopterons ici la proposition de Höder (2012 : 245) qui nous invite à étendre ce concept aux différentes langues et à prolonger donc cette notion de diasystème en la transformant en cadre pour l'étude d'une multitude de situations d'acquisition et/ou apprentissage d'une langue autre que celle (ou celles) que l'on considère comme L1 du sujet. Le processus principal qui conduit à réellement établir un diasystème, c'est-à-dire un qui serait intégré dans la cognition du locuteur et dans l'usage, est l'identification inter-linguistique (Weinreich 1964) fondée sur la relation de similarité entre les éléments des différentes variétés, perçus comme équivalents par les locuteurs multilingues selon des critères structurels, phoniques, sémantiques, pragmatiques et autres, souvent eux-mêmes en compétition (Höder 2012 : 244). La similarité typologique entre les langues est, selon toute évidence, un facteur central pour l'établissement du diasystème, tout comme la notion de l'usage, puisque, comme le souligne Ellis (2013 : 365), « *if the units of language are constructions, then language acquisition is the learning of constructions. So L2A depends upon learners' experience of language usage and upon what they can make of it* ».

Cela pose, inévitablement, la question du rôle à accorder à la sémantique et aux représentations liées à celle-ci. Comment en pourrait-il être autrement, lorsque le lexique bilingue associe, selon l'expression de Grosjean (2001) *one mind two languages*, un même contenu sémantique (et/ou conceptuel) à deux formes totalement différentes du point de vue formel ? Comme le fait déjà remarquer Grosjean (1998) l'usage des termes « sémantique » et « conceptuel », de ce qui distingue ces deux termes et de ce qui les rapproche, n'est pas toujours clair dans la bibliographie relative au bilinguisme, qui, autour des années 2000, correspondait à un domaine nouveau. Si l'on considère que le terme « conceptuel » recouvre des traits et des aspects relevant de l'extralinguistique, on peut se demander si cette surface conceptuelle correspond à la L1 ou à la L2. À notre avis, cette question terminologique persiste jusqu'à aujourd'hui, 25 ans plus tard, et serait liée au peu d'intérêt que suscitent, dans les approches mettant l'accent sur l'interlangue et plus généralement sur le diasystème, les facteurs d'ordre sémantique et conceptuel. Cela est particulièrement le cas pour les langues « autres », c'est-à-dire celles qui présentent des caractéristiques formelles et morphologiques les éloignant, à des degrés divers, de l'anglais et des langues occidentales les plus étudiées. L'une de lacunes qu'ambitionne de combler ce numéro est précisément celle-ci, à la fois au niveau de la description linguistique que du fonctionnement du sujet, locuteur natif et/ou apprenant d'une langue. De ce point de vue, la notion du transfert s'avère porter un intérêt particulier, qu'elle relève du purement linguistique ou du culturel. On peut définir sommairement le transfert comme une situation dans laquelle des

analogies entre les items de deux (ou plusieurs) langues faisant partie du répertoire du sujet sont repérées et exploitées par ce dernier, sans que cela ne corresponde nécessairement à des processus conscients. À partir de cette définition, l'on peut formuler des interrogations qui consistent à distinguer ce qui relève du phénomène aléatoire basé sur l'analogie de ce qui fait partie de mécanismes caractérisés par des degrés (variables) de systématité et de régularité, sur le plan phonologique, morpho-sémantique ou encore extralinguistique et culturel.

Concernant le premier niveau d'analyse linguistique, la contribution de Xiaoliang Luo porte sur la phonologie du chinois et l'enseignement de la prononciation. L'auteur observe que l'enseignement de la prononciation, reconnu comme l'un des principaux défis de l'enseignement du chinois langue étrangère, demeure relativement peu exploré sur les plans empirique et théorique (Klötter 2015). Xiaoliang Luo souligne par ailleurs l'importance de s'intéresser à la phonologie plutôt qu'à la seule phonétique, car « l'acquisition de la prononciation ne consiste pas à maîtriser des sons isolés, mais bien à intégrer des oppositions phonémiques ». L'auteur mobilise les notions de surdité phonologique et de crible phonologique pour cibler les difficultés de prononciation rencontrées par les apprenants francophones, directement liées aux traits distinctifs présents en chinois mais absents en français. Il associe une analyse détaillée des différences phonologiques entre le chinois et le français à des propositions de correction phonétique fondées sur les traits distinctifs du chinois, susceptibles d'aider les apprenants à surmonter le crible perceptif induit par leur langue maternelle. L'approche proposée repose sur un entraînement perceptif précédant la production et s'inscrivant dans un processus à long terme, en lien étroit avec l'acquisition lexicale.

La perception constitue l'objet de diverses contributions de ce volume, dont plusieurs concernent la morphologie : celle d'Hélène Giraudo, Serena Dal Maso et Sabrina Piccinin, porte sur la conscience morphologique, définie comme l'habileté à identifier et à manipuler des morphèmes, les plus petites unités de sens dans les mots, et qui constitue un facteur essentiel pour l'efficacité en lecture. Cette connaissance métalinguistique est, avec la conscience phonologique, l'un des corrélats les plus forts de réussite dans le processus de lecture et de compréhension des mots. L'objectif de cette contribution est d'examiner le rôle de différentes dimensions de la conscience morphologique pour l'efficacité en lecture chez des élèves de fin CE1 (7-8 ans) scolarisés en REP (réseau d'éducation prioritaire). L'étude qui est présentée ici rapporte les résultats obtenus auprès d'un échantillon de participants qui a effectué plusieurs épreuves et plus précisément : deux questionnaires, sept tests standardisés, quatre épreuves morphologiques et une de compréhension de texte. Un accent particulier est mis sur la morphologie dérivationnelle du français qui représente un champ d'investigation privilégié pour l'étude des processus de création des mots, en raison de sa variation morphologique et sa complexité.

La dérivation est le domaine sur lequel porte la contribution de Nikolaos Ntagkas, qui présente une étude expérimentale menée avec le protocole d'amorçage non masqué, dans laquelle la perception et le traitement de plusieurs types de mots, construits et non construits, sont examinés. L'analyse morphologique est sémantiquement contrainte, dans la mesure où les morphèmes sont d'abord identifiés sur la base de leur sens, ensuite sur la base de leur forme. Ainsi, selon l'hypothèse de l'auteur, les mots qui sont sémantiquement transparents sont traités comme des mots construits, ayant une structure morphologique interne ; ceux qui sont semi-transparentes sont traités comme des mots construits, à condition que leur sens puisse être connecté à celui des morphèmes qui les constituent, au travers de processus sémantiques (ex. métaphore, métonymie, etc.), sur le plan synchronique ; les mots sémantiquement opaques, de l'autre côté, sont non-construits. Les deux expériences, effectuées auprès de locuteurs natifs du grec, leur méthodologie et leurs résultats, après analyse statistique, sont présentés en détail. L'interprétation des résultats, en rapport aux différentes approches et modèles de traitement morphologique, insiste sur le fait qu'il convient d'envisager le morphème comme une représentation d'ordre morphosémantique, non assimilable à des représentations situées avant le niveau des représentations lexicales.

Toujours dans le domaine de la morphologie, la contribution d'Assimakis Filatouras, Madeleine Voga et Anna Anastassiadis-Syméonidis, concerne l'étymologie populaire qu'elle envisage comme un processus de nature psycholinguistique. Il s'agit d'un processus universel en même temps qu'un processus historique, étant donné qu'il se manifeste à plusieurs périodes d'une langue, en l'occurrence du grec. Son départ peut être non-conscient, lorsqu'il prend appui sur la difficulté de lier de manière simultanée le signifiant au signifié d'une unité lexicale d'un point de vue phonologique et morphologique : un nouveau rapprochement est ainsi tenté. Le départ peut être également conscient, à caractère ludique ou normatif. Une grande quantité de données linguistiques est présentée, fournissant des indices en faveur de l'idée qu'un certain degré de systématisme et de régularité, non seulement au niveau phonologique mais aussi morphosémantique, existe dans l'étymologie populaire du grec. L'action phonologique, allant du degré zéro au niveau syllabique et pluri-morphémique, ainsi que les règles phonologiques susceptibles de jouer, sont présentées, de même que les quatre niveaux d'action morphologique. Le lien avec l'expérience du locuteur, liée non seulement à son expérience de sujet parlant mais aussi à celle du monde qui l'entoure, ainsi que sa contribution à instaurer ou à renforcer la transparence de l'unité lexicale faisant l'objet d'un rapprochement parétymologique, sont soulignés.

La contribution d'Alberto Bramati traite du démonstratif « ça » et de son usage, dans une perspective contrastive français-italien. Selon l'auteur, ce démonstratif, longtemps interprété comme une simple variante familière et orale de « cela », présente en réalité des propriétés syntaxiques et sémantiques distinctes, ce qui rend sa description et sa traduction dans d'autres langues particulièrement complexes.

Cette étude se fonde sur un corpus bilingue composé de 24 textes français contemporains en prose (9 romans, 11 pièces de théâtre et 4 essais en sciences humaines) et de leurs traductions en italien. À partir de ce corpus, l'auteur choisit de se concentrer sur « ça » sujet, en raison de la forte fréquence de cette fonction dans le corpus bilingue (45 % des 2 968 occurrences du pronom). L'auteur identifie cinq fonctions du pronom « ça » (déictique, anaphorique intraphrastique, anaphorique interphrastique, indéfini, sujet disloqué) et analyse les solutions adoptées par les traducteurs professionnels pour rendre cette forme en italien. Les données traductologiques recueillies montrent que le pronom neutre « ciò », seul démonstratif italien monosyllabique, apparaît rarement, en raison de la différence de registre par rapport au pronom français « ça », tandis que les deux autres pronoms démonstratifs italiens (« questo » et « quello ») sont principalement utilisés lorsque le pronom français reprend une proposition. Dans les autres cas, d'autres solutions sont privilégiées, telles que l'omission du pronom, la lexicalisation ou le recours à des structures comme [è (una cosa/qualcosa) che].

La contribution de Jing Guo se situe au croisement de la didactique et de la pragmatique dans une perspective inter-langues. L'autrice y analyse les représentations de la politesse chinoise chez les apprenants francophones ainsi que les difficultés qu'ils rencontrent dans son apprentissage. Elle s'appuie sur un questionnaire portant sur divers aspects de la politesse, telles que les termes d'adresse, les actes de salutation et les actes de remerciement, distribué aux étudiants de différents cours de chinois L2 à l'Institut nationale des langues et civilisations orientales (INALCO). À partir des données recueillies, Jing Guo examine les phénomènes de transfert culturel dans la perception des comportements polis et impolis, ainsi que les phénomènes de filtrage culturel, qui amènent l'apprenant à interpréter les informations d'une langue étrangère et les comportements de ses locuteurs à travers le prisme de sa propre langue et de sa culture maternelles. L'autrice analyse les stades de l'adaptation culturelle, par lesquels l'apprenant ajuste progressivement sa perception et ses comportements à la culture de la langue cible. Les données recueillies permettent par ailleurs de dégager des pistes de recherche, notamment concernant le sentiment d'identité contradictoire chez les apprenants francophones d'origine chinoise. Jing Guo esquisse également des pistes didactiques, en s'appuyant sur le traitement de la politesse dans le manuel *Méthode de chinois*, où les codes de politesse sont intégrés dans les différentes unités, et en évoquant des activités orales mobilisant à la fois les compétences sociolinguistiques, linguistiques et pragmatiques.

Le fonctionnement du locuteur et les processus cognitifs qui sous-tendent le traitement langagier en langue seconde et/ou étrangère constituent l'objet des deux derniers articles de ce numéro. Plus précisément, le lexique mental bilingue (grec-français) constitue le thème de la contribution de Madeleine Voga, qui prend comme point de départ le constat suivant : malgré les efforts effectués depuis environ dix ans, les données relatives à la perception et au traitement bilingues

peinent à s'intégrer dans une approche qui pourrait permettre de considérer, sous un angle unificateur, une variété de situations expérimentales concernant la proximité formelle et morphologique des mots des langues étudiées. L'article examine un certain nombre de travaux expérimentaux, tous menés grâce au protocole chronométrique de l'amorçage masqué, qui, en l'occurrence, est inter-langue et inter-alphabet, puisqu'il concerne la combinaison de langues grec-français. Les principes d'une approche qui vise à unifier les données, d'une part intra-alphabet et d'autre part inter-alphabet, en particulier grec-français, sont présentés. Cette approche se fonde, d'une part, sur les transferts d'activation (co-activation des langues) dans le lexique mental, et d'autre part sur le rôle accru attribué à la représentation morphologique. Le propos est illustré par une revue de données chronométriques issues de protocoles d'amorçage masqué grec-français, dans la direction L1 vers L2, dans lesquels différents types de cognats et non-cognats sont testés. Ce type de protocole expérimental permet l'analyse fine des flux d'activation au sein du lexique mental, dont le caractère constructionnel et la structuration morphosémantique sont soulignés.

Enfin, la contribution de Enhao Zheng rapporte les données d'une expérience avec le protocole de mouvements oculaires, employé dans très peu de recherches en chinois L2, ce qui confère à cette étude un caractère novateur. Cette étude s'intéresse à l'acquisition du vocabulaire par la lecture, et plus particulièrement à l'attention accordée par l'apprenant francophone de niveau intermédiaire aux mots inconnus contenus dans des textes. Les résultats, présentés à partir des différentes mesures que l'oculométrie met en place (durée de la première fixation, nombre de visites et de revisites oculaires, temps total passé sur un mot donné) montrent que les apprenants n'accordent pas la même attention à tous les mots inconnus rencontrés dans le texte, *i. e.*, à tous les mots-cibles. Diverses variables entrent en jeu — notamment la présence, dans le mot, de caractères comportant des composants familiers, une forme atypique, ou encore la position du mot dans le texte, qui conduisent à des pistes pour faciliter l'apprentissage lexical.

Pour conclure, les contributions réunies dans ce numéro s'intéressent à la fois au lexique et à la grammaire des langues, en mobilisant une grande variété d'approches et en explorant de multiples dimensions, notamment phonologiques, morphologiques et pragmatiques. Les auteurs s'appuient sur des corpus linguistiques, des retours d'expérience, des questionnaires, ainsi que sur des protocoles expérimentaux. Ces démarches permettent d'appréhender les compétences lexicales et grammaticales sous différents angles, tout en enrichissant les ressources disponibles sur des langues et des combinaisons linguistiques pour lesquelles la recherche, qu'elle soit théorique, appliquée ou didactique, demeure encore relativement récente ou limitée.

Comme le suggère le sous-titre de ce numéro, l'objectif est de contribuer à la mise à disposition de données sur ces langues, mais aussi d'informations, d'observations et d'analyses portant sur les usages linguistiques. Les contributions

examinent également les perceptions des langues et les représentations que s'en font les locuteurs et les apprenants, analysées selon diverses perspectives théoriques et approches méthodologiques. Dans ce cadre, les études présentées dans ce numéro abordent des enjeux essentiels liés à l'apprentissage linguistique. Outre les défis de l'acquisition lexicale et grammaticale, elles s'intéressent aux différentes formes de filtrage associées à la langue et à la culture maternelles, ainsi qu'aux phénomènes de transfert qui se manifestent tant sur le plan linguistique que culturel. Les analyses montrent comment ces facteurs influencent la manière dont les locuteurs natifs et non natifs et les apprenants d'une langue étrangère perçoivent, interprètent et produisent les formes linguistiques dans la langue cible. Enfin, une partie des contributions esquisse des pistes pour l'enseignement/apprentissage des langues, en mettant en lumière des stratégies susceptibles de favoriser l'acquisition du lexique, de la grammaire et des usages pragmatiques dans des contextes variés.

Références bibliographiques

- BOOIJ, Gert, *The morphology of Dutch*, Oxford: Oxford University Press, 2002.
- GROSJEAN, François, « Studying bilinguals: Methodological and conceptual issues », *Bilingualism: Language and Cognition*, 1 (1998): p. 131-149.
- GROSJEAN, François, « The bilingual's language modes », dans J. NICOL (éd.), *One Mind, Two Languages: Bilingual Language Processing*, Oxford: Blackwell, 2001, p. 1-22.
- ELLIS Nick, « Form, Function, and Frequency in L1 and L2 Learning of Constructions », dans T. HOFFMANN & G. TROUSDALE (éd.), *The Oxford Handbook of Construction Grammar*, Oxford: Oxford University Press, 2013, p. 365-378.
- HÖDER, Stephen, « Multilingual constructions. A diasystematic approach to common structures », dans K. BRAUNMULLER & C. GABRIEL (éd.), *Multilingual individuals and multilingual societies*, Amsterdam/Philadelphie: John Benjamins, 2012, p. 241-258.
- KLÖTER Henning, « Chinese as a Foreign Language, Linguistics and Pedagogy », dans R. SYBESMA (dir.), *Encyclopedia of Chinese Language and Linguistics Online*. Brill. 2015: https://doi.org/10.1163/2210-7363_ecll_COM_000184.
- JACKENDOFF, Ray et Jenny AUDRING, *The texture of the lexicon. Relational morphology and the parallel architecture*, Oxford: Oxford University Press, 2020.
- KIPARSKY, Paul, « Lexical Phonology and Morphology », dans I.S. YANG & The Linguistic Society of Korea (éd.), *Linguistics in the morning calm*. Seoul: Hanshin, 1982, p. 3-91.

- LINN, Andrew, Danielle CANDEL, Jacqueline LEON, « Présentation : Linguistique appliquée et disciplinarisation », *Histoire Épistémologie Langage*, 33 (1) « *Linguistique appliquée et disciplinarisation* » (2011) : p. 7-14.
- RASTIER, François « La sémantique cognitive. Éléments d'histoire et d'épistémologie », dans B. NERLICH (éd.), *Histoire Épistémologie Langage* 15 (1), « *Histoire de la sémantique* », (1993) : p. 153-187.
- STEGU, Martin, « La linguistique appliquée : discipline ou groupement de disciplines indépendantes? », *Histoire Épistémologie Langage* 33 (1) « Linguistique appliquée et disciplinarisation » (2011) : p. 129-139.
- WEINREICH, Uriel, « Is a Structural Dialectology Possible? », *Word* 10 (1954) : p. 388-400.
- WEINREICH, Uriel, *Languages in Contact: Findings and Problems*. 3rd printing. London: Mouton [New York 1953], 1964.

Enseignement de la prononciation du chinois aux étudiants francophones : une méthode basée sur la phonologie

Xiaoliang Luo

*Université Paris Cité/
Histoire des Théories Linguistiques, UMR7597*

Introduction

L'enseignement de la prononciation constitue à la fois la partie la plus rudimentaire et un des défis majeurs de l'enseignement du chinois langue étrangère (CLE). Paradoxalement, les réflexions théoriques sur la méthode de l'enseigner restent rares. Dans le contexte français, Zhang-Colin & Gianninoto (2022) traitent brièvement de l'influence du structuralisme après-guerre dans l'enseignement de la phonétique ; Allanic (2023), dans son article centré sur l'enseignement des sinogrammes, évoque dans quelques passages la question du *pinyin* et la dissociation oral/écrit. Dans le contexte chinois, ces dernières années, plusieurs études s'intéressent à l'enseignement de la prononciation en CLE. Wáng Lǐjiā 王理嘉 (2005), dans une des rares études théoriques, analyse en profondeur l'influence du *pinyin* sur l'apprentissage de la prononciation, en formulant des recommandations pédagogiques. Dèng Dān 邓丹 (2020) applique la phonétique dans la pratique pédagogique. Huáng Dānlì 黄丹丽 (2011), après avoir examiné plus d'une centaine d'articles, résume les tendances récentes de la recherche sur l'enseignement de la phonétique du chinois langue étrangère (CLE) en trois caractéristiques principales : 1) un nombre croissant d'études analysent les difficultés concrètes rencontrées par les apprenants ; 2) l'introduction de la phonétique expérimentale a contribué à rendre la recherche sur l'enseignement de la prononciation plus scientifique ; 3) les études ciblant des groupes d'apprenants selon leur langue maternelle se multiplient également.

Pour ce qui concerne les apprenants francophones, Liú & Qí 刘苏乔、齐冲 (2004) identifient, à partir d'un corpus constitué d'enregistrements d'apprenants

français, les difficultés principales de ce public dans l'acquisition des consonnes. Gāo & Shí 高玉娟、石峰 (2006) adoptent une approche expérimentale pour examiner les effets de transfert positif et négatif de la langue maternelle sur l'apprentissage des voyelles chinoises. Ces travaux ont beaucoup contribué à l'enseignement de la prononciation du chinois, tant au niveau théorique qu'au niveau pratique.

Notre travail s'inscrit dans les courants structuralistes d'après-guerre appliqués à l'enseignement des langues. Les objectifs de cette étude sont les suivants : 1) de clarifier les questions théoriques, à savoir ce que l'on apprend réellement lorsque l'on apprend la prononciation d'une langue seconde ; 2) de proposer une pratique d'enseignement opérationnelle et exhaustive, adaptée à un public francophone.

Dans la seconde moitié du xx^e siècle, dans le cadre de la phonologie structurale, l'école dite de « Saint-Cloud/Zagreb », représentée par des chercheurs tels que Guberina (Croatie) et Rivenc (France), a formulé des idées novatrices sur l'acquisition de la phonétique en langue seconde. Selon cette école, l'échec dans la prononciation d'une langue étrangère s'explique par le système phonologique de la langue maternelle, qui filtre automatiquement les sons n'existant pas phonologiquement dans ce système (cf. Guberina, 1965ab ; Rivenc, 2002, 2003). Autrement dit, si une opposition phonologique n'existe pas dans la langue maternelle, l'apprenant aura du mal à la percevoir et donc à la produire. Bref, ce qui ne peut être perçu ne peut être produit.

S'appuyant sur cette hypothèse, nous analyserons les difficultés typiques rencontrées par les apprenants francophones lors de l'apprentissage du chinois, largement observées en contexte universitaire français. En rupture avec l'approche qui enseigne des sons isolés, cette étude développe une stratégie de correction phonétique fondée sur les traits distinctifs du chinois, afin d'aider les apprenants à surmonter le crible perceptif induit de leur langue maternelle. Nous discuterons d'abord les fondements théoriques de notre approche et résumerons ensuite les interférences du *pinyin*¹, dans l'apprentissage de la prononciation et en analysons les causes. Nous aborderons de manière exhaustive les difficultés typiques rencontrées par les francophones dans le domaine du consonantisme et du vocalisme, avant d'examiner l'accent et le rythme dans l'apprentissage de la prononciation du chinois, des aspects encore peu étudiés dans les travaux antérieurs.

Fondements théoriques

Le concept de « surdité phonologique » a été proposé pour la première fois par Polivanov (1931), linguiste faisant partie du Cercle de Prague, afin de décrire

1. Influencé par les systèmes de romanisation existants, le *pinyin* a été élaboré collectivement par des linguistes chinois dans les années 1950. Approuvé en 1958 par l'Assemblée nationale populaire à Pékin comme système officiel de romanisation, il est depuis devenu la norme internationale pour la transcription du chinois et constitue aujourd'hui un outil incontournable dans l'enseignement du chinois langue étrangère (CLE).

les obstacles systématiques rencontrés par les apprenants qui n'entendent pas les sons absents dans leur langue maternelle. Cette notion a été confirmée par des recherches contemporaines, notamment dans le domaine de la neurolinguistique. Ainsi, Vaissière (2020) cite des études montrant que, bien que les nourrissons exposés au chinois standard soient initialement capables de distinguer les consonnes voisées des non voisées, cette capacité se perd à l'âge adulte, précisément parce que cette opposition n'est pas pertinente dans le système phonologique de cette langue.

Ce concept de « surdité phonologique » a été ensuite développé par Troubetzkoy (1939) en « crible phonologique », qui explicite une défaillance structurelle chez les apprenants :

Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phonèmes. [...] Chaque homme s'habitue dès l'enfance à analyser ainsi ce qui est dit et cette analyse se fait d'une façon tout à fait automatique et inconsciente. Mais en outre le système des cribles, qui rend cette analyse possible, est construit différemment dans chaque langue. L'homme s'approprie le système de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend le crible phonologique de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue étrangère entendue, il se produit de nombreuses erreurs et incompréhensions. Les sons de la langue étrangère reçoivent une interprétation phonologiquement inexacte, puisqu'on les fait passer par le crible phonologique de sa propre langue. (Troubetzkoy 1976, 54)

Sous l'influence du Cercle de Prague, Guberina (1965a, 1965b) met en évidence une telle observation empirique : chez les apprenants adultes, même lorsque l'appareil auditif fonctionne parfaitement, certaines distinctions phonologiques échappent totalement à la perception. Ces erreurs ne relèvent donc pas d'une incapacité articulatoire ni d'une défaillance de l'appareil auditif, mais d'un problème perceptif lié au crible phonologique. Cette hypothèse a des implications pédagogiques majeures et affirme que toute correction phonétique efficace doit d'abord traiter les erreurs de perception avant de viser la production¹.

Partant de cette hypothèse, une question centrale se pose : que faut-il faire entendre aux apprenants pendant l'enseignement de la langue seconde ? Selon Jakobson (1976, 85), « Ce qui importe [...], ce n'est pas du tout l'individualité [...] de chacun d'eux, vue en elle-même et existant pour elle-même. Ce qui importe, c'est leur opposition réciproque au sein d'un système [...]. » Ce ne sont donc pas des sons isolés ni des phonèmes pris individuellement qu'il faut faire travailler les apprenants, mais les traits distinctifs au sein d'un système qui permettent la compréhension et la production correctes d'une langue. Cette perspective constitue

1. Pour une histoire du développement de la méthode Saint-Cloud/Zagreb en phonétique corrective, cf. RIVENC (2003).

une rupture nette avec les approches qui consistent à faire répéter mécaniquement des sons sans les inscrire dans un système d'oppositions.

Prenons l'exemple d'un francophone apprenant le chinois. Bien que la langue française possède des sons aspirés, comme [p^h], cette aspiration n'a aucune valeur distinctive : les sons [p] et [p^h] sont deux variantes du même phonème /p/. En conséquence, lorsqu'un francophone entend un mot chinois qui commence par une consonne aspirée, comme pǎo [p^hau] 跑, il ne perçoit pas l'aspiration, filtrée par le crible phonologique du français. Dans cet exemple, il est inutile de demander à l'apprenant de répéter mécaniquement des consonnes aspirées ; une stratégie plus efficace consiste à lui présenter des paires minimales en chinois, comme bǎo/ pǎo, afin de l'aider à percevoir la différence entre les traits aspiré/non-aspiré qui distinguent deux mots différents.

Ainsi, la stratégie pédagogique proposée repose sur trois objectifs progressifs : d'abord, amener l'apprenant à percevoir les traits distinctifs du chinois à travers des exercices de discrimination avec des paires minimales ou des quasi-paires minimales ; ensuite, lui faire comprendre la fonction distinctive de ces traits, dont la confusion crée immédiatement une confusion sémantique ; enfin, travailler la production une fois que la perception est stabilisée, en recourant des méthodes adaptées aux francophones.

Quant à l'introduction du *pinyin* indispensable pour l'apprentissage du chinois, elle ne devrait intervenir qu'à un stade ultérieur, une fois les bases perceptives solidement acquises, afin de minimiser l'influence négative du *pinyin* sur l'apprentissage de la prononciation.

Les difficultés posées par le *pinyin*

Wáng Lǐjiā 王理嘉 (2005) montre de manière convaincante que le système de transcription chinois, communément appelé « *pinyin* » (cf. note 1), repose non pas sur une correspondance directe entre symboles et sons, mais sur la représentation phonémique, autrement dit, contrairement à l'API qui suit (ou au moins prétend suivre) le principe d'un symbole = un son, le *pinyin* est un système phonémisant. Ce système, conçu pour les locuteurs natifs dans les années 1950, est particulièrement efficace pour annoter la prononciation des caractères chinois. Toutefois, il introduit un certain nombre de difficultés et d'obstacles pour les apprenants étrangers, notamment en raison des relations complexes entre symboles graphiques et sons, des variations contextuelles de prononciation, ainsi que des formes abrégées. Les recommandations de Wáng Lǐjiā aux enseignants de CLE sur comment enseigner le *pinyin* négligent toutefois selon nous deux aspects importants : sur le plan théorique, elles négligent les différences fondamentales du processus d'acquisition chez les locuteurs natifs et chez les apprenants langue seconde ; sur le plan pratique, ces recommandations manquent de faisabilité.

Sur le plan théorique, un enfant sinophone natif et un apprenant de chinois langue seconde n'abordent pas le *pinyin* de la même manière. L'enfant sinophone, ayant déjà intériorisé le système phonologique de sa langue maternelle avant d'être confronté au *pinyin* à l'école, apprendra le *pinyin* comme un outil d'annotation, qui ne peut en aucun cas interférer avec la perception ni la production. À l'inverse, les apprenants étrangers dépendent d'un système de transcription pour apprendre le chinois, avant que les catégories phonémiques ne soient bien établies. Ils doivent donc construire leur système phonologique du chinois langue seconde en même temps qu'ils apprennent le *pinyin* (et non pas à l'aide de ce dernier puisque son influence est souvent négative). Or, le caractère phonémique abstrait de ce système peut perturber leur apprentissage : dans le système *pinyin*, une même lettre peut représenter plusieurs sons selon le contexte, et différents graphèmes peuvent représenter un même son. Cela crée une charge cognitive importante pour les apprenants, qui peinent à prononcer correctement un mot en lisant la transcription *pinyin*, et qui peuvent donner par erreur une valeur phonétique de l'orthographe française à un symbole du *pinyin* en confondant les deux, puisqu'ils ont souvent des difficultés pour comprendre qu'un symbole n'a de valeur phonétique que dans un système orthographique donné, au moins pendant une période plus ou moins longue de l'apprentissage.

Sur le plan pratique, les recommandations de Wáng Lǐjiā (2005) présupposent que les enseignants sont capables de maîtriser les principes fondamentaux de la phonétique et de la phonologie, de ne pas confondre les deux, d'expliquer de manière précise les distinctions entre phonèmes, variantes et phones, ainsi que de clarifier les correspondances souvent très complexes entre les symboles du *pinyin* et les sons prononcés. En réalité, peu d'enseignants possèdent à la fois ces connaissances théoriques et la capacité de les transposer efficacement dans une approche pédagogique adaptée aux débutants.

De ce fait, bien que le travail de Wáng Lǐjiā constitue une étape importante dans la recherche sur l'enseignement de la prononciation du CLE sur le plan théorique, et qu'il soit très éclairant pour un enseignant idéal disposant d'une solide formation en phonétique et en phonologie, il reste difficile à mettre en œuvre dans un contexte d'enseignement où la majorité des enseignants ne sont pas formés en phonétique, et encore moins en phonologie.

Face à ces limites, la présente étude propose une approche alternative, fondée sur un entraînement perceptif précédant la production. L'idée est de ne pas introduire immédiatement le *pinyin*, mais de commencer par des exercices ciblés permettant aux apprenants francophones de développer une sensibilité aux traits distinctifs du chinois qui distinguent les mots sur le plan sémantique. Ce n'est qu'une fois cette perception stabilisée que le *pinyin* peut être introduit, dans le but de minimiser les interférences et de consolider une représentation phonologique cohérente avant l'introduction du *pinyin*. Bien entendu, ce processus d'apprentissage ne se réalise pas en une seule fois, mais doit être répété de manière régulière,

chaque répétition incluant la perception et la distinction des sons avant l'apprentissage du *pinyin*.

Traits distinctifs absents chez les francophones
et stratégies d'enseignement

Selon le concept de crible phonologique, la principale difficulté pour les apprenants francophones réside dans le fait que les sons inexistants en français sont automatiquement filtrés au niveau perceptif. Les travaux précédents, notamment Liu & Qi (2004), Gao & Shi (2006), montrent que les difficultés d'apprentissage observées relèvent précisément de ce type de phones. Comme cela a été établi précédemment, ce que les apprenants doivent acquérir, ce ne sont pas des phonèmes isolés, mais les traits distinctifs au sein du système phonologique. La présente étude dresse une liste systémique de toutes les oppositions phonémiques du chinois absentes en français (voir tableau 1), et propose pour chacune d'elles des stratégies de correction phonétique spécifiques.

Tableau 1

	Traits distinctifs (combinabilité possible)
Consonnes	Aspiration
	Affrication
	Alvéolopalatalité
	Rétroflexion
	Vélarité
Voyelles	Voyelle de type <i>schwa</i>
	Diphthongues orales
	Diphthongues nasales
Tons	Hauteur (<i>pitch</i>)

Les tons étant toujours au cœur des discussions dans des études de CLE, de nombreuses études antérieures y sont consacrées, nous ne nous y attarderons donc pas davantage.

Dans les universités françaises, les manuels largement utilisés actuellement, tels que *Méthode de chinois — Premier niveau* (Rabut *et al.*, 2017), adoptent une structure pédagogique où la phonétique est traitée de manière intensive dès les premières leçons: avant même le début de l'apprentissage de la langue, deux leçons préparatoires sont consacrées au *pinyin* (et aux caractères), après lesquelles aucun enseignement systématique de la phonétique n'est poursuivi. Le manuel *Le Chinois comme en Chine* (Allanic, 2023) suit un schéma similaire, concentrant l'apprentissage du *pinyin* et de la prononciation dans les trois premières leçons.

Selon nous, cette structure didactique reflète une conception implicite selon laquelle la phonétique serait une compétence secondaire et accessoire : les étudiants pourraient en assimiler les bases en très peu de temps, puis passer sans transition au « contenu sérieux », à savoir les caractères chinois, la grammaire, la compréhension orale et écrite, ainsi que l'expression.

La présente étude défend une position opposée : cette organisation, dominante dans les manuels de chinois en usage en France, occulte complètement le caractère progressif de l'acquisition phonétique. Elle trahit une négligence manifeste de la part des concepteurs de manuels à l'égard de l'enseignement de la prononciation. Or, les autres compétences linguistiques — compréhension et expression orales, compréhension et expression écrites, sinogrammes — sont toutes abordées selon une logique de progression à long terme. Pourquoi en irait-il autrement pour la phonétique ? Peut-on sérieusement penser que l'apprentissage de la prononciation en chinois peut être mené à bien en seulement deux ou trois semaines ?

Une heureuse exception est le manuel *Ni shuo ne* (Arslangul *et al.*, 2009), dans lequel la phonétique et la prononciation s'étalent sur le volume en entier, mais il est souvent utilisé dans le secondaire et rarement dans le supérieur, notamment dans les filières LEA et LLCER.

S'appuyant sur les travaux du Cercle de Prague — notamment ceux de Polivanov, Troubetzkoy et Jakobson — la présente étude soutient que l'élément le plus important dans l'apprentissage de la prononciation est l'acquisition des traits distinctifs, car ce sont eux qui assurent la fonction distinctive et permettent de différencier les mots. Dès lors, des entraînements phonétiques intensifs et de courte durée, réalisés en tout début d'apprentissage, alors que les apprenants ne disposent pratiquement d'aucun bagage lexical, les placent dans une situation de confusion : répéter et imiter des mots dont ils ignorent la signification ne leur permet pas de comprendre ce qu'ils sont censés apprendre sur le plan phonétophonologique. Ce type de formation, déconnectée de la différenciation sémantique, produit des résultats médiocres au prix d'un effort considérable.

À l'inverse, la présente étude considère que l'apprentissage de la prononciation doit s'inscrire dans un processus à long terme, en lien étroit avec l'acquisition lexicale. C'est dans l'identification progressive de la signification des mots et des expressions que les apprenants peuvent intérioriser les traits distinctifs pertinents du système phonologique du chinois. Ce n'est qu'en comprenant comment ces traits permettent de distinguer les unités de sens que l'apprenant peut en percevoir l'importance et les intégrer réellement. L'objectif n'est donc pas de répéter mécaniquement, mais de comprendre comment la langue organise la signification avec le système phonologique.

Dans notre pratique d'enseignement, la progression pédagogique repose sur une distinction explicite entre deux phases de l'apprentissage : la phase de perception, qui précède celle de production. Avant de pouvoir prononcer correctement, l'apprenant doit d'abord être en mesure de percevoir avec précision les traits distinctifs

du chinois. Pour cela, un travail intensif d'écoute ciblée est nécessaire, l'objectif étant de « savoir entendre » avant de « savoir prononcer ». Ces deux phases ne sont pas linéaires, mais s'entrelacent tout au long de l'apprentissage, avec à chaque fois la perception en amont de la production.

Dans ce cadre, deux types d'exercices peuvent être mis en place pour développer les capacités perceptives des apprenants. Le premier type consiste à discriminer deux sons entendus et à décider s'ils sont identiques ou différents. Le second type demande aux apprenants d'identifier un mot entendu parmi plusieurs termes au choix phonétiquement proches. Les mots utilisés dans ces exercices doivent faire partie du lexique déjà appris ou en cours d'acquisition, et sont choisis pour leur proximité phonétique. Les apprenants doivent donc s'appuyer sur les traits distinctifs spécifiques pour les différencier de manière précise. À titre d'exemple, des paires telles que 知道 *zhīdào* / 迟到 *chídào* peuvent être utilisées pour travailler l'opposition entre les sons rétroflexes aspirés et non aspirés ; 假期 *jiàqī* / 夏季 *xiàjì* pour la distinction affriquées/fricatives ; 来到 *láidào* / 累倒 *lèidào* pour la différence de hauteur entre diphtongues, etc. Ces mots sont en général rencontrés par les apprenants dès le niveau débutant et avant le niveau intermédiaire. La capacité de percevoir les traits distinctifs est une condition préalable à la bonne compréhension et à la maîtrise de ce lexique. Sans cette perception fine, les apprenants ne peuvent accéder ni à une prononciation correcte, ni à une compréhension précise de la signification.

Dans ce qui suit, nous examinerons, suivant l'ordre établi dans le tableau 1, les stratégies de correction phonétique des traits distinctifs du chinois qui posent des difficultés aux apprenants francophones.

Les consonnes

L'aspiration

En chinois, l'aspiration constitue un trait distinctif impliquant plusieurs séries de consonnes. Les consonnes aspirées et non aspirées (par exemple *p* et *p^h*) permettent de distinguer de nombreux mots. À l'inverse, en français, bien que des sons aspirés puissent exister phonétiquement, ils n'ont pas de valeur distinctive. L'aspiration a surtout une fonction expressive au sens troubetzkoyen, telle que la colère, l'étonnement, etc. Par conséquent, les francophones ont des difficultés à percevoir la différence entre aspirées et non-aspirées.

Afin d'aider les apprenants francophones à acquérir l'opposition phonémique entre consonnes aspirées et non-aspirées, l'enseignement peut débuter par des exercices de discrimination, avec des paires minimales citées plus haut comme 饱 *bǎo* et 跑 *pǎo*. Une fois que les étudiants parviennent à percevoir leur différence, la phase de production peut être abordée suivant trois pistes. Premièrement, une conscientisation de l'aspiration est nécessaire. Bien que l'aspiration ne soit pas un

trait distinctif en français, elle assume une fonction expressive, comme l'a défini Troubetzkoy. On peut ainsi s'appuyer sur des exemples tirés du français — par exemple l'exclamation « Imp[p^h]ossible ! » — pour aider les étudiants à prendre conscience qu'ils sont tout à fait capables de produire des consonnes aspirées dans certains contextes. Deuxièmement, il est possible de mobiliser l'anglais comme langue intermédiaire. En anglais, les consonnes sourdes en position initiale et en syllabe accentuée sont aspirées (par exemple dans « paper », « top »). Comme l'anglais constitue souvent la première langue vivante des apprenants francophones, cette familiarité est susceptible de faciliter la maîtrise des consonnes aspirées du chinois. Troisièmement, l'enseignant peut exploiter des environnements phonétiques favorables à la production de l'aspiration en chinois, en particulier les contextes accentués. À titre d'exemple, dans la phrase « 我特别累 ! » (*Wǒ tèbié lèi*), l'accent de phrase porte sur la syllabe 特 (*tè*), ce qui favorise la production d'une consonne aspirée. La maîtrise de l'aspiration dans ce type de contexte accentué permet aux étudiants de prendre appui sur une production réussie, qu'ils peuvent ensuite généraliser à d'autres mots et contextes.

Les affriquées

Les occlusives et les affriquées — comme celles représentées en *pinyin* par les graphèmes <d> et <z> — ne sont pas distinctives en français, alors qu'elles le sont en chinois, ce qui rend leur distinction indispensable pour éviter toute confusion. Bien que l'opposition affriquées/fricatives — comme celle entre <z> et <s> en *pinyin* — assume également une fonction distinctive, les apprenants francophones tendent davantage à confondre les occlusives et les affriquées, plutôt que les affriquées et les fricatives. Cette tendance s'explique par le fait que le français possède une opposition bien établie entre occlusives et fricatives, mais ne reconnaît pas l'affriquée comme une catégorie phonémique à part entière. Ainsi, lorsqu'un locuteur francophone entend une affriquée comme [ts], il a tendance à la filtrer automatiquement comme l'occlusive [t], et non comme une fricative [s].

L'enseignement doit, là encore, débiter par des exercices d'écoute ciblés, recourant à des paires minimales ou quasi-minimales telles que 他 *tā* / 擦 *cā* ou 现在 *xiànzài* / 现代 *xiàndài*, afin de développer la capacité des apprenants à distinguer ces deux séries de consonnes. Une fois cette perception établie, trois pistes peuvent être expérimentées pour aider les francophones à produire correctement les affriquées. Premièrement, il convient de s'appuyer sur la langue française pour sensibiliser les apprenants à la notion d'affriquée, bien que celle-ci ne soit pas phonémique en français. En effet, les affriquées existent en français phonétiquement, en phonotactique, comme dans l'enchaînement des mots « ôte son » de la phrase « Il ôte son manteau », où le contact entre [t] et [s] crée une articulation du type [ts] affriqué. Ce type de phénomène, perceptible dans la chaîne parlée, peut servir de point d'appui pour construire une conscience articuloire. On observe par

ailleurs la présence d'affriquées en français dans certaines variétés régionales (au Québec par exemple) ou sociolectales (récupérée d'ailleurs par l'extrême droite qui prétend un lien entre affrication et Afrique, n'étant pas suffisamment cultivée pour avoir entendu parler de l'affrication en gallo-roman). Deuxièmement, certains emprunts lexicaux en français incluent des affriquées, comme le mot « tsar », qui peut être utilisé pour illustrer concrètement ce type de son. Troisièmement, l'anglais comme langue intermédiaire présente un grand nombre d'affriquées, phonémique ou phonotactique. Une fois que les étudiants ont pris conscience de cette catégorie de sons, les apprenants francophones disposant d'une bonne base en anglais sont généralement capables d'en transférer la maîtrise vers le chinois, ce qui facilite grandement l'apprentissage de ces sons.

Les alvéolo-palatales

La distribution des consonnes alvéolo-palatales (*j*, *q*, *x* en *pinyin*) en chinois est relativement complexe : elles présentent une certaine complémentarité partielle avec les vélaires (occlusives et fricative), rétroflexes (affriquées et fricative) et dentales (affriquées et fricative), mais elles entrent en opposition phonologique avec les occlusives dentales, comme l'illustrent les paires minimales 记 *jì* et 弟 *dì*, ou 起 *qǐ* et 体 *tǐ*. Or, c'est précisément cette opposition que les apprenants francophones ont tendance à confondre. En effet, chez une partie des francophones, les consonnes dentales [t] et [d] sont à la fois affriquées et palatalisées devant les voyelles antérieures. Par exemple, le mot « tirer » peut être prononcé [tiʁe] ou [tʃiʁe], comme deux variantes libres. Pour les francophones, les alvéolo-palatales seront automatiquement filtrées comme des occlusives dentales une fois passées par le crible phonologique. Leur distinction avec les dentales est donc difficilement perceptible pour eux.

Au cours de l'enseignement, suite à des exercices d'écoute ciblés, une fois que les étudiants sont capables de percevoir la différence entre les alvéolo-palatales et les dentales, l'enseignant peut aider la production en commençant par une conscientisation. Il s'agira notamment de leur faire prendre conscience des alvéolo-palatales comme variation sociologique du français, dans certains milieux défavorisés. L'enseignant peut ensuite les encourager à tenter de produire ces sons dans des environnements phonétiques favorables, comme devant la voyelle *i*, qui facilite naturellement la palatalisation. Par ailleurs, il n'est pas nécessaire d'exiger une maîtrise immédiate et parfaite de la prononciation de ces sons. Pour les apprenants francophones ayant un bon niveau d'anglais, il peut être utile, en cas de difficulté articulatoire, de recourir aux sons palataux de l'anglais, tels que [tʃ] ou [tʃʰ], comme étapes intermédiaires dans l'apprentissage de la prononciation chinoise.

Les rétroflexes

Les rétroflexes ne possèdent pas non plus de statut phonémique en français. Une minorité de locuteurs natifs francophones peuvent produire les rétroflexes comme des variantes libres des fricatives palatales, comme dans le mot « chat » ([ʃa]), parfois prononcé [ʂa]. En revanche, en chinois, les rétroflexes et les dentales (*z, c s* en *pinyin*) s'opposent phonologiquement.

Si les chinois natifs confondent assez souvent rétroflexes et dentales, les francophones évitent en général aisément cette confusion. Cela s'explique par l'opposition des palatales et dentales en français, que les francophones ne confondent que très rarement. Étant donné que les fricatives palatales et les rétroflexes présentent des caractéristiques acoustiques proches, certains apprenants francophones tendent à substituer les palatales aux rétroflexes. Cela pourrait sembler en contradiction avec un autre phénomène mentionné précédemment : une partie des apprenants francophones ont également tendance à remplacer les alvéolo-palatales en chinois (comme *j, q, x*) par des palatales ([tʃ], [tʃʰ], [ʃ]) issues du français ou d'autres langues.

En réalité, cette confusion des alvéolo-palatales et des rétroflexes n'a pas d'effet critique sur l'expression orale en chinois, car ces deux séries de sons ne forment aucune opposition phonologique en chinois. Leur répartition est en effet complémentaire : les alvéolo-palatales apparaissent exclusivement devant des voyelles palatales (comme [i], [y]) ou des semi-voyelles correspondantes, tandis que les rétroflexes n'apparaissent jamais dans ces contextes.

Dans la pratique pédagogique, l'enseignant peut commencer par expliquer les principes articulatoires des sons rétroflexes, en invitant les étudiants à essayer de replier la pointe de la langue vers l'arrière. Toutefois, une méthode plus efficace consiste à associer les rétroflexes à des voyelles vélaires, c'est-à-dire à des voyelles postérieures, comme dans les mots 书 *shū*, 住 *zhù*, ou 出 *chū*. Le contexte articulatoire créé par ces voyelles postérieures facilite la production des rétroflexes. Une fois ces sons relativement maîtrisés dans ce contexte favorable, les apprenants peuvent généraliser cette production à d'autres combinaisons consonne-voyelle, y compris celles qui ne contiennent pas de voyelle postérieure. En réalité, cette influence des voyelles sur la prononciation des fricatives palatale est observable en français. Par exemple, la fricative palatale [ʃ] est bien produite comme une vraie palatale lorsqu'elle est suivie d'une voyelle antérieure, comme dans « chic » ou « chute ». Mais lorsqu'elle est combinée avec la voyelle [u], comme dans le mot « chou », les locuteurs natifs du français, sous l'effet de la voyelle postérieure, ont souvent tendance à produire spontanément une rétroflexe.

La fricative vélaire [x] (notée h en *pinyin*)

En français, il existe des occlusives vélaires [g], [k] alors que la fricative vélaire est absente. Cela rend la fricative vélaire du chinois (<h> en *pinyin* et [x] en API)

difficile à percevoir et à produire pour les francophones. À cette difficulté phonologique s'ajoute l'influence de l'orthographe française, dans laquelle la lettre <h> est muette, ce qui retarde une prise de conscience chez eux du fait que le <h> du *pinyin* doit impérativement être prononcé. C'est pourquoi nous recommandons, avant même d'introduire la lettre <h> dans l'apprentissage du *pinyin*, de travailler au préalable la perception et la production du son correspondant, afin de réduire au maximum l'influence négative du système graphique sur l'acquisition de ce son.

La fricative vélaire est absente en français, phonologiquement et phonétiquement. Il est donc utile, en classe, de commencer par des comparaisons avec les occlusives vélares du chinois, [k] et [kʰ], afin de faire percevoir [x] par contraste occlusive/fricative et insister sur leur lieu d'articulation identique.

Comme le son [x] présente des caractéristiques acoustiques proches de celles de la fricative glottale [h], les apprenants rencontrant des difficultés peuvent utiliser le [h] de l'anglais comme prononciation transitoire dans un premier temps, car cela n'entraînera aucune confusion en chinois : en chinois, il n'existe pas d'opposition phonémique entre [x] et [h], d'autant que [x] peut parfois subir un processus de débuccalisation, c'est-à-dire la perte de son lieu d'articulation dans la cavité buccale pour se réaliser comme une simple fricative glottale [h]. Par exemple, dans le mot 快活 *kuàihuo*, le second caractère 活 est en position atone, il subit un affaiblissement prosodique, et le [x] de *huó* tend à se réduire vers [h] (voire à disparaître presque complètement pour conduire à une syllabe [wo] au lieu de [xwo]). Encore une fois, cette variation phonétique n'affecte pas la compréhension, car elle ne crée pas d'ambiguïté sémantique dans le système phonologique du chinois.

Ordre de difficulté et progression

Dans les parties précédentes, une analyse systémique des difficultés rencontrées par les locuteurs francophones dans l'apprentissage des consonnes du chinois a été proposée. Ces difficultés sont directement liées aux traits distinctifs présents en chinois mais absents en français. Toutefois, certaines consonnes chinoises cumulent plusieurs de ces traits distinctifs, ce qui rend leur acquisition plus complexe. Ainsi, nous estimons que l'apprentissage du système consonantique chinois par les francophones devrait suivre une progression structurée. Le tableau 2 présente le degré relatif de complexité des consonnes du chinois pour les apprenants francophones : plus une consonne comporte de traits distinctifs à acquérir (indiqués par un astérisque), plus elle est considérée comme relativement complexe. Dans la pratique pédagogique, il est recommandé de guider les apprenants dans un apprentissage allant du plus simple au plus complexe, en commençant par les consonnes présentant un seul trait distinctif absent en français, pour introduire progressivement celles qui en cumulent deux ou trois. Prenons l'exemple de la série rétroflexe <zh>, <ch> et <sh> : le son <sh> ne présente que le trait de rétroflexion absent en français ; <zh> y ajoute le trait d'affrication ; <ch> combine enfin rétroflexion, affrication et

aspiration. Par conséquent, l'ordre d'apprentissage des rétroflexes devrait être le suivant : <sh>, <zh>, <ch>.

Tableau 2

Consonnes	Labiales	Dentales	Alvéolo-palatales	Rétroflexes	Vélaires
Occlusives	b[p] p[p ^h]*	d[t] t[t ^h]*			g[k] k[k ^h]*
Affriquées		z[ts]* c[ts ^h]**	j[tɕ]** q[tɕ ^h ***]	zh[tʂ]** ch[tʂ ^h ***]	
Fricatives	f[f]	s[s]	x[ɕ]*	sh[ʂ]*	h[x]*
Nasales	m[m]	n[n]			
Liquides		l[l]			
Approximantes				r[ʮ]**	

Les voyelles

Les voyelles de type schwa

En chinois, deux voyelles sont fréquemment filtrées par le « crible phonologique » des francophones et automatiquement assimilées à une voyelle centrale moyenne, également appelée schwa [ə]. Il s'agit, d'une part, de la voyelle [ɤ], comme dans la syllabe *zhè* (« 这 »), et d'autre part, de la voyelle [ɨ], comme dans la syllabe *zhī* (« 只 »).

La première, [ɤ], est une voyelle postérieure non arrondie, ayant à peu près le même degré d'ouverture qu'une voyelle moyenne. Or, le français ne possède pas de voyelle postérieure non arrondie, et [ɤ] présente des caractéristiques acoustiques proches de celles d'une voyelle centrale. Elle est donc facilement assimilée au schwa par les francophones. La seconde, [ɨ], ne diffère du schwa que par son degré d'ouverture, ce qui la rend elle aussi sujette à la confusion et à une perception erronée comme une voyelle centrale.

Pour les apprenants francophones, la distinction entre ces deux sons est en général difficile à établir. Il est donc nécessaire de recourir à de nombreux exercices d'écoute ciblés fondés sur des paires minimales, telles que 舌 *shé* / 十 *shí*, 这 *zhè* / 只 *zhǐ*, 车 *chē* / 尺 *chǐ*, 色 *sè* / 四 *sì*, afin d'aiguiser leur perception auditive. Ces exercices peuvent ensuite être étendus à des formes polysyllabiques, comme 时候 *shíhou* / 舌头 *shétou*, même si ces exemples ne sont pas tous des paires minimales *stricto sensu*, ils constituent néanmoins un support efficace pour développer la sensibilité auditive des apprenants à ces oppositions vocaliques.

Dans un second temps, une fois qu'une certaine compétence en perception est établie, l'enseignant peut passer à une approche articulatoire, en mobilisant les principes de la phonétique articulatoire. Étant donné que les mouvements des lèvres sont visibles, l'enseignant peut les démontrer aux apprenants. Pour produire [ɤ], il peut partir du schwa français, et guider les étudiants à étirer les lèvres vers

les côtés, afin de supprimer l'arrondissement et obtenir une voyelle postérieure non arrondie. Pour produire [i], il peut procéder de la même manière, en partant également du schwa, mais en fermant les lèvres, afin de produire une voyelle centrale mais fermée, le [i].

Les diphtongues orales

En français, il n'existe pas de diphtongues. Les francophones, lors de l'apprentissage des diphtongues du chinois, ont souvent tendance à transférer les diphtongues de l'anglais. Or, les diphtongues anglaises et chinoises présentent des différences phonétiques importantes, et l'application des diphtongues anglaises peut entraîner une confusion entre les paires <ai/ei> et <ao/ou> en chinois.

Plus précisément, les diphtongues anglaises et chinoises diffèrent sur deux points : premièrement, dans les diphtongues anglaises [aɪ] et [aʊ], le segment [a] est une voyelle postérieure, tandis que dans les diphtongues chinoises [ai] et [au], le [a] est une voyelle antérieure ; deuxièmement, la durée de transition entre le timbre ouvert et le timbre fermé est nettement différente entre les deux langues, Lindau *et al.* (1990) montre que cette durée est beaucoup plus longue en anglais qu'en chinois. C'est précisément pour cette raison que les diphtongues anglaises sont plus faciles à distinguer pour les francophones, tandis que celles du chinois, avec leur transition plus brève et moins marquée, sont plus difficiles à différencier. Quand les apprenants francophones calquent sur l'anglais, ils s'attendent à une différence très marquée entre <ai/ei> et entre <ao/ou> et peinent à percevoir cette différence moins marquée en chinois.

Dans la pratique pédagogique, l'enseignant doit concevoir des exercices d'écoute ciblés pour entraîner les apprenants à entendre les différences entre les deux termes au sein de chacune des deux paires de diphtongues. Par exemple, les oppositions 高 *gāo* / 狗 *gǒu* et 来 *lái* / 楼 *lóu* peuvent être utilisées pour renforcer la perception auditive. Une fois que les étudiants parviennent à distinguer ces sons, l'enseignant peut procéder à des démonstrations articulatoires, en montrant visuellement les mouvements de la bouche et le temps de transition entre les deux timbres de la diphtongue, afin que les apprenants puissent les reproduire correctement.

Les diphtongues nasales

En français, on trouve à la fois des monophthongues nasales (par exemple, <en>, <on> en orthographe française) et des séquences de voyelle + consonne nasale (VN, par exemple, des mots comme « bonne » [bɔ̃n], « fine » [fin]). En revanche, les finales nasales en chinois sont constituées exclusivement de combinaisons de type VN (par exemple, <in>, <an> en *pinyin*). Pour les francophones, ces finales nasales posent principalement trois types de difficultés.

Premièrement, ils ont tendance à prononcer les séquences VN comme des voyelles nasales, en calquant sur le français, où la nasalisation affecte directement la voyelle pour former une monophthongue. Deuxièmement, ils confondent facilement les nasales antérieures et postérieures, c'est-à-dire les oppositions <an/ang> et <en/eng>. Il convient de noter que l'opposition <in/ing> a déjà disparu chez une grande partie des locuteurs natifs du chinois. Troisièmement, dans les séquences VN du français, la consonne nasale N est réalisée comme une occlusive nasale avec trois phases complète pour une occlusive : fermeture > tenue > relâchement, alors qu'en chinois, la consonne nasale en fin de syllabe n'est pas relâchée, il n'y a que deux phases : fermeture et tenue. Les francophones ont donc souvent tendance à relâcher la nasale finale en prononçant le chinois, ce qui entraîne une erreur articuloire et une confusion sémantique, par exemple, en relâchant la nasale finale dans 看 *kàn*, le mot sera perçu par un locuteur chinois comme une séquence 看呢 *kàn-ne* avec [n] relâché perçu comme une particule modale en fin de phrase.

Pour faire la distinction entre nasales antérieures et postérieures, les apprenants doivent être exposés à un grand nombre d'exercices de discrimination auditive, avec des paires comme 反向 *fǎnxiàng* / 方向 *fāngxiàng* ou 一分 *yīfēn* / 一封 *yīfēng*, afin de renforcer et stabiliser la perception de leur différence. Au stade de la production, il est nécessaire d'insister sur le fait que la finale VN en chinois correspond à l'enchaînement de deux sons successifs (voyelle + nasale), et non à une monophthongue nasale comme en français.

L'enseignant peut également s'appuyer sur des emprunts à l'anglais présents en français, comme « parking », pour éveiller chez les apprenants une conscience articuloire de la nasale vélaire [ŋ]. Enfin, qu'il s'agisse de nasales antérieures ou postérieures, il est essentiel de corriger la tendance à relâcher la consonne finale, en montrant les gestes articuloires en chinois, qui s'observent visuellement, s'arrêtant à la phase de tenue, sans continuer sur la phase de relâchement.

Accentuation et rythme

Bien que le chinois soit une langue à tons, la proportion élevée de mots dissyllabiques et polysyllabiques dans le chinois contemporain fait que la fonction distinctive des tons a un rôle moins déterminant. Selon nous, si l'apprentissage des tons demeure important, une prononciation imparfaite des tons n'entrave pas nécessairement la compréhension du sens et la perfection des tons peut attendre un niveau plus avancé. Par ailleurs, les réserves à la méthode actuelle d'enseigner des tons se font entendre assez souvent chez les enseignants de CLE, qui en même temps peinent à trouver une meilleure solution. À titre d'exemple, le 3^e ton est enseigné, dans tous les manuels de CLE à notre connaissance, comme un ton descendant-montant, alors qu'en réalité, il est soit prononcé uniquement dans

sa première partie, soit modifié par la règle du sandhi tonal¹. Il est donc rare qu'il apparaisse sous sa forme complète décrite dans les manuels — avec une courbe descendante puis montante — ce qui crée une grande confusion chez les apprenants après l'apprentissage théorique de ce ton.

Par rapport aux tons, l'accentuation et le rythme jouent un rôle plus décisif dans l'expression orale du chinois.

Concernant l'accentuation, bien que l'existence d'accent phonologique soit très controversée dans les études théoriques, il est incontestable que certaines oppositions lexicales ou phrastiques dépendent de la place de l'accentuation. Par exemple, les mots 神话 (*shénhuà*, « mythe ») et 神化 (*shénhuà*, « diviniser ») se distinguent non par les segments ni par les tons, mais par leur schéma accentuel : le premier suit une structure iambique, le second une structure trochaïque (puisque le causatif 化 *-huà* est un suffixe qui ne porte pas de proéminence prosodique). Par ailleurs, une accentuation existe aussi dans les unités lexicales plurisyllabiques fortement fusionnées : dans les mots composés comme 知道 *zhīdao*, 朋友 *péngyou*, ou 前面 *qiánmian*, l'accentuation de la première syllabe conduit souvent à l'affaiblissement de la syllabe finale, qui devient un ton neutre dans la plupart des cas.

Au niveau phrastique, les mots grammaticaux et les pronoms sont en général non accentués, tandis que les mots pleins et les mots porteurs de sémantique comportent souvent une syllabe accentuée. Par exemple, dans une phrase tirée d'un manuel de chinois pour débutants : 她女朋友叫他别喝酒 (*Tā nǚpéngyou jiào tā bié hē jiǔ* « Sa petite amie lui a dit de ne pas boire d'alcool »), l'accentuation tombe sur les mots 朋 *péng*, 叫 *jiào* et 别 *bié*.

En outre, les phrases en chinois s'organisent autour de ces syllabes accentuées pour former des groupes rythmiques et des pauses. Dans l'exemple ci-dessus, la phrase peut être divisée en trois groupes rythmiques : *Tā nǚpéngyou* 她女朋友 / *jiào tā* 叫他 / *bié hē jiǔ* 别喝酒.

Par rapport à ces réalités phonétiques de la langue parlée — omniprésence de l'accentuation et du rythme — l'enseignement phonétique du chinois accuse un retard considérable. Ni les caractères chinois, ni le *pinyin* ne reflètent les accents réels de l'oral. Les manuels de chinois ne signalent pas non plus les syllabes accentuées dans les phrases, ni ne proposent de segmentation rythmique. Cela crée une réelle difficulté pour les apprenants dans la progression de leur expression orale.

Ce problème est particulièrement marqué chez les francophones. Par rapport à d'autres langues européennes comme l'anglais, l'allemand, l'italien ou l'espagnol, le français se distingue par l'absence d'accent phonologique : l'énergie est répartie de manière relativement égale sur toutes les syllabes, sauf la dernière qui est accentuée. Les francophones perçoivent difficilement les accents et ont tendance

1. Deux troisièmes tons (noté T₃) consécutifs donnent lieu au changement du premier T₃, qui est conventionnellement considéré comme devenir un deuxième ton (noté T₂). Cette règle est formalisée comme suit : T₃+T₃ > T₂+T₃. Par exemple, dans 你好 *nǐhǎo*, 你 *nǐ* portant le T₃ sera réalisé avec un contour montant.

à ignorer spontanément ceux présents dans la langue cible. Cela se manifeste de manière visible dans les stéréotypes véhiculés dans les œuvres audiovisuelles, où les francophones, de façon caricaturale, parlent anglais en effaçant quasiment les accents de l'anglais.

Dans ce contexte, il est nécessaire que l'enseignant, dans sa pratique pédagogique, attire explicitement l'attention des apprenants francophones sur la position des accentuations en chinois et les aide à identifier ces proéminences prosodiques dans les phrases, à les marquer et à segmenter les énoncés en groupes rythmiques. Si les francophones produisent un discours sans alternance entre syllabes fortes et faibles, accentuée et non-accentuée, sans faire de pauses entre groupes rythmiques, en distribuant l'énergie uniformément et à vitesse constante sur chaque syllabe, cela impose un effort cognitif important à l'auditeur chinois pour comprendre l'énoncé. Il est donc indispensable que les enseignants accompagnent les apprenants dans la construction d'une sensibilité à l'accentuation, à l'emphase et au rythme dans la langue chinoise.

Conclusion

Cet article examine, du point de vue de la phonologie à la place de la phonétique, les difficultés rencontrées par les francophones dans l'apprentissage de la phonétique du chinois, ainsi que les origines de ces difficultés. En s'appuyant sur la pratique pédagogique, il propose un ensemble structuré de dispositifs pédagogiques et de stratégies de correction phonétique adaptés à ce public.

Par rapport aux travaux précédents, cette étude se distingue principalement par deux innovations. D'une part, elle applique au domaine de l'acquisition d'une langue seconde le concept de « crible phonologique » troubetzkoyen, en mettant l'accent sur le principe pédagogique selon lequel la perception précède la production. D'autre part, elle défend l'idée que l'acquisition de la prononciation ne consiste pas à maîtriser des phonèmes isolés, mais à percevoir les différences entre les sons, c'est-à-dire à développer une compétence fondée sur les traits distinctifs au sein du système phonologique de la langue cible.

Nous espérons que cette contribution, modeste dans son ambition, pourra ouvrir de nouvelles perspectives théoriques dans le domaine de la recherche sur la phonétique du CLE, tout en offrant aux enseignants de terrain des outils concrets et opérationnels pour leur pratique pédagogique.

Références bibliographiques

ALLANIC, Bernard, *Théorie et pratique de l'enseignement raisonné des sinogrammes — Contribution à la didactique du chinois langue étrangère*, Mémoire d'HDR, université Paul-Valéry Montpellier 3, 2023.

- ARSLANGUL, Arnaud, Claude LAMOUROUX et Isabelle PILLET, *Ni Shuo Ne?*, Paris: Didier, 2009.
- DENG Dan 邓丹, « Hanyu yuyan yanjiu yu duiwai hanyu yuyin jiaoxue 汉语语音研究与对外汉语语音教学 » [Études sur la phonétique du chinois et enseignement de la phonétique du chinois langue étrangère], *Guoji hanyu jiaoyu* 国际汉语教育 [Éducation internationale du chinois], 5.2 (2020): p. 51-59.
- GAO Yujuan 高玉娟 et Feng SHI 石峰, « Faguo xuesheng hanyu yuanyin xuexi zhong de muyu qianyi de shiyan yanjiu 法国学生汉语元音学习中母语迁移的实验研究 » [Études expérimentales du transfert de la langue maternelle dans l'apprentissage des voyelles du chinois par les apprenants français], *Yunnan shifan daxue xuebao (duiwai hanyu jiaoxue yu yanjiu ban)* 云南师范大学学报 (对外汉语教学与研究版) [Bulletin de l'université Normale de Yunnan, Enseignement et recherches du chinois langue étrangère], 2006, 4: p. 33-36.
- GUBERINA, Petar, « La méthode structuro-globale audio-visuelle », *Revue de Linguistique Appliquée*, 1 (1965a): p. 11-30.
- GUBERINA, Petar, « La parole dans la méthode structuro-globale-audio-visuelle », *Le Français dans le Monde*, 103 (1965b): p. 49-54.
- HUANG Danli 黄丹丽, « Duiwai hanyu yuyin jiaoxue yanjiu zongshu 对外汉语语音教学研究综述 » [Synthèse de la recherche sur l'enseignement de la phonétique du chinois langue étrangère], *Kejiao wenhui* 科教文汇 [Recueil en sciences et éducation], 13 (2011): p. 75-77.
- JAKOBSON, Roman, *Six leçons sur le son et le sens*, Paris: Minuit, 1976.
- LINDAU, Mona, Kjell NORLIN et Jan-Olof SVANTESSON, « Some cross-linguistic differences in diphthongs », *Journal of the International Phonetic Association* 20.1 (1990): p. 10-14.
- LIU Suqiao 刘苏乔 et Chong QI 齐冲, « Faguo xuesheng xuexi hanyu fuyin zhong de yixie wenti 法国学生学习汉语辅音中的一些问题 » [Quelques problèmes dans l'apprentissage des consonnes du chinois chez les étudiants français], *Yuyan wenzi yingyong* 语言文字应用 [Application de langue et écriture], 2004.4: p. 76-81.
- POLIVANOV, Evgenij D., « La perception des sons d'une langue étrangère », *Travaux du Cercle Linguistique de Prague*, 4 (1931): p. 79-96.
- RABUT, Isabelle, Hong LIU et Zhitang DROCOURT, *Méthode de chinois premier niveau*, 3^e édition, Paris: L'Asiathèque, 2017.
- RIVENC, Paul, « Place et rôle de la phonétique dans la méthodologie SGAV », dans R. RENARD (dir.), *Apprentissage d'une langue étrangère / seconde, vol. 2, La phonétique verbo-tonale*. Louvain-la-Neuve: De Boeck Supérieur, 2002, p. 25-34.

- RIVENC, Paul, « Brève histoire de la Problématique SGAV, Étapes dans la construction d'une méthodologie », dans P. RIVENC (dir.), *Apprentissage d'une langue étrangère/seconde, vol. 3, La méthodologie*. Louvain-la-Neuve : De Boeck Supérieur, 2003, p. 85-134.
- TRUBETZKOY, Nikolaï S., *Grundzüge der phonologie*, Prague, 1939.
- TRUBETZKOY, Nikolaï S., *Principes de phonologie*, traduit par Jean Cantineau. Paris : Klincksieck, [1949] 1976.
- VAISSIÈRE, Jacqueline, *La phonétique*. Paris : Presses universitaires de France, 2020.
- WANG Lijia 王理嘉, « “Hanyu pinyin fangan” yu shijie hanyu yuyin jiaoxue 《汉语拼音方案》与世界汉语语音教学 », *Shijie hanyu jiaoxue 世界汉语教学* [Enseignement du chinois global], 2, (2005) : p. 5-11.
- ZHANG-COLIN, Ying & Mariarosaria GIANNINOTO, « Les idées didactiques en chinois langue étrangère (CLE) depuis 1945 : recoupements et croisements avec l'histoire du français langue étrangère (FLE) », *Documents pour l'histoire du français langue étrangère et seconde*, 68 (2022) : p. 1-12.

Conscience morphologique et lecture chez les élèves en Réseau d'Éducation Prioritaire : une étude pilote en fin de CE1

Hélène GIRAUDO*, Serena DAL MASO** & Sabrina PICCININ**

*Laboratoire CLLE, CNRS & Université de Toulouse Jean-Jaurès

**Université de Vérone, Vérone, Italie

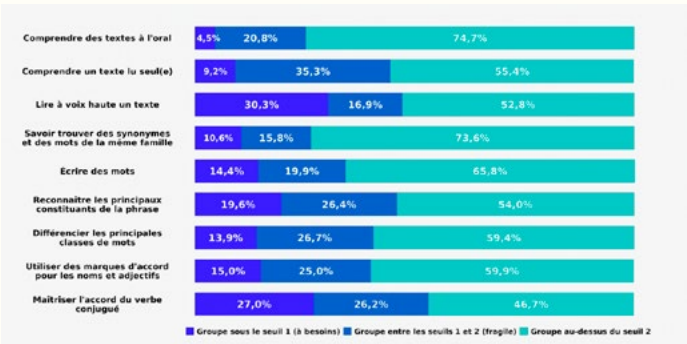
La littéracie comme enjeu de développement économique et social et de lutte contre les inégalités

La notion de littéracie est définie par le dictionnaire *Le Robert* comme « Aptitude à lire, à comprendre et à utiliser l'information écrite dans la vie quotidienne. » Selon cette définition, même si l'école en assume la responsabilité première à travers sa mission de formation, l'acquisition de la littéracie dépasse le seul cadre scolaire. Elle concerne donc non seulement les résultats des élèves, mais aussi leur capacité à fonctionner efficacement dans la société, des enjeux plus larges de développement économique et social promus par les pays de l'OCDE. Selon cette organisation, la littéracie, à côté de la numératie, fait partie des « socles essentiels à l'acquisition de connaissances et compétences plus poussées, dans un monde où l'apprentissage passe en grande partie par la compréhension de l'écrit et le traitement d'informations écrites [...]. Une maîtrise de base de ces compétences est indispensable pour s'instruire, se former, participer au marché du travail et à la vie sociale et civique » (OCDE, 2024: 28, 48). La dernière *Évaluation des compétences des adultes* réalisée en 2024 et à laquelle ont participé 31 pays, nous alerte sur le recul des compétences en littéracie de la grande majorité des pays¹, dont la France (*i.e.*, -6.70 points en 10 ans pour la France). Elle nous informe par ailleurs que ces reculs sont particulièrement marqués dans les segments les moins instruits de la population. En effet, les écarts entre les compétences moyennes

1. Sur les dix dernières années, seuls le Danemark et la Finlande ont connu une hausse significative des compétences en littéracie chez les adultes.

en littéracie dépendent du niveau d’instruction des parents des adultes évalués dans l’étude : Ceux ayant des parents très instruits ont un niveau de compétences supérieur à ceux dont les parents sont moins instruits (*e.g.*, l’écart ajusté entre les scores en littéracie est de 23 points en France), pointant ainsi un effet du statut socio-économique des individus sur leurs compétences langagières. Ces résultats suggèrent que l’école ne parvient plus à compenser les inégalités initiales liées au milieu socio-économique, en laissant ainsi persister les écarts de compétences en littéracie entre les individus selon leur origine sociale.

Parallèlement aux enquêtes internationales de type PISA, qui mesurent l’efficacité des systèmes éducatifs, le ministère de l’Éducation nationale (MEN) en France organise des évaluations standardisées des élèves à différents niveaux de leur scolarité¹. L’évaluation des acquis en français et en mathématiques des élèves, scolarisés du CP au lycée (*i.e.*, de 6 à 18 ans), est pratiquée annuellement afin d’obtenir des indicateurs pour mesurer l’efficacité du système éducatif. Les dernières évaluations nationales de 2024 chez les élèves de CM2 (*i.e.*, Cours Moyen 2, soit la dernière année de l’école primaire où les enfants sont âgés de 10 à 11 ans), révèlent qu’un nombre important d’entre eux sont en difficulté lorsque sont mesurées leurs compétences en français (Cf. graphique 1).



Graphique 1 — Répartition des élèves dans les groupes selon la compétence évaluée en français en début de CM2 en septembre 2024. Données issues du document de travail n° 2024-E14 du MEN, p. 22.

Les difficultés des groupes identifiés comme ayant des besoins de remédiation ou fragiles (seuils 1 et 2), se manifestent aux 3 niveaux du traitement impliqués dans la lecture, selon les proportions d’effectifs suivantes :

- au niveau formel (*i.e.*, « lire un texte à haute voix » et « écrire des mots ») : 40,75 %,
- au niveau syntaxique (*i.e.*, « reconnaître les principaux constituants de la

1. <https://www.education.gouv.fr/l-evaluation-des-acquis-des-eleves-du-cp-au-lycee-1208g>.

- phrase », « différencier les principales classes de mots », « utiliser des marques d'accord pour les noms et les adjectifs », « maîtriser l'accord du verbe conjugué »): 22,47 %,
- au niveau sémantique (*i.e.*, « comprendre des textes à l'oral », « comprendre un texte lu seul », « savoir trouver des synonymes et des mots de la même famille »): 16,3 %.

Lorsque sont comparés les résultats des élèves scolarisés dans des établissements des secteurs REP et REP+ (*i.e.* Réseau d'Éducation prioritaire, un réseau d'établissements où les efforts éducatifs sont intensifiés pour prévenir l'échec scolaire, où sont scolarisés 20,3 % du total des élèves en France à la rentrée 2023¹), à ceux des élèves en dehors de ces secteurs, des écarts de l'ordre de 16 à 25 points sont constatés sur les différentes compétences évaluées en français, ce qui montre que ces élèves ont globalement une maîtrise plus faible de la langue. Le rapport précise que l'écart le plus important (-25,8 points) concerne la compréhension d'un texte lu seul(e) (p. 26). Ces données suggèrent qu'en dépit d'une politique éducative spécifique mise en place en France depuis plus de 40 ans et qui vise à « corriger l'inégalité [sociale] par le renforcement sélectif de l'action éducative dans les zones et dans les milieux sociaux où le taux d'échec scolaire est le plus élevé » (circulaire du 9 juillet 1981), la littéracie des élèves à profil fragile demeure inférieure à celle des autres élèves, lesquels ont déjà, selon les données de 2023, des compétences langagières en recul de -6,70 points par rapport à 2013 (OCDE, 2024). Or, le référentiel de l'éducation prioritaire établit comme priorité numéro 1, de « garantir l'acquisition du “lire, écrire, parler” et enseigner plus explicitement les compétences que l'école requiert pour assurer la maîtrise du socle commun² ».

L'impact du profil sociolinguistique des élèves sur l'acquisition de la littéracie

Si la lecture, apprentissage complexe intégrant divers processus et connaissances (Grabe & Stoller, 2019 ; Castles *et al.*, 2018), se fait sans effort chez les lecteurs adultes expérimentés, elle constitue un véritable défi pour les enfants, en particulier pour ceux ayant un profil fragile. Les fragilités résultent souvent de l'exposition des enfants à une langue peu riche et peu variée, qui ne comprend pas par exemple les formes écrites, comme chez les enfants qui ne sont pas impliqués dans des activités de lecture partagée avec leurs parents, phénomène fréquent dans les milieux

1. Deux types de réseaux sont identifiés par le MEN : les REP+ qui concernent les quartiers ou les secteurs isolés connaissant les plus grandes concentrations de difficultés sociales ayant des incidences fortes sur la réussite scolaire et les REP, plus mixtes socialement mais rencontrant des difficultés sociales plus significatives que celles des collèges et écoles situés hors de l'éducation prioritaire.

2. éducol : <https://eduscol.education.fr/1028/la-politique-de-l-education-prioritaire-les-reseaux-d-education-prioritaire-rep-et-rep>.

socio-économiques faibles ou défavorisés (*e.g.*, Dias-Broens & van Steensel, 2022). Cette problématique se cumule également avec celle des enfants bilingues parlant une langue minoritaire à la maison (EBLM) avec des répercussions potentiellement néfastes à long terme sur leurs performances académiques dans la langue de scolarisation (*e.g.*, Kieffer, 2010). Les travaux de littérature consacrés aux effets du statut économique et social des familles et/ou d'une langue maternelle minoritaire sur l'acquisition de la littéracie ont montré que ces deux facteurs, souvent positivement corrélés, touchent conjointement le développement langagier des apprenants (*e.g.*, Bellocchi & Bonifacci, 2024 ; Bonifacci *et al.*, 2020 ; 2022 ; Calvo & Bialystok, 2014 ; Casalis, Deacon & Pacton, 2011). Les élèves issus des catégories sociales défavorisées ont des compétences langagières moins bonnes que ceux des catégories moyennes et supérieures, et les bilingues de langue minoritaire obtiennent de moins bons résultats que les monolingues. Les travaux de Bellocchi et Bonifacci (2024) ont exploré les effets différenciés du SES et du bilinguisme chez des enfants scolarisés en CE2 (*i.e.*, 8-9 ans), CM1 (*i.e.*, 9-10 ans) et CM2 (*i.e.*, 10-11 ans) dans une école non-REP en France. Quatre groupes d'élèves (*i.e.*, monolingues SES bas/moyen/haut et bilingues SES bas/moyen/haut) ont été testés sur leur compétences non-verbales (*i.e.*, intelligence non-verbale) et verbales orales (*i.e.*, connaissance verbale et compréhension morpho-syntaxique, mémoire phonologique à court terme), leurs compétences en lecture (*i.e.*, fluidité en lecture de texte, lecture de mots réguliers et irréguliers et de pseudo-mots) et en écriture (*i.e.*, écriture de mots réguliers et irréguliers et de pseudo-mots). Globalement, leurs résultats indiquent un effet négatif du Statut Économique et Social (SES), qui impacte de manière large les habiletés langagières ainsi que certaines compétences cognitives, telles que le raisonnement non verbal et la mémoire phonologique. Par ailleurs, le type particulier de bilinguisme de ces élèves (*i.e.*, les enfants bilingues parlant une langue minoritaire) est associé à des difficultés plus marquées dans les composantes verbales de la littéracie, notamment les connaissances lexicales et morpho-syntaxiques, la fluence en lecture de texte ainsi que la précision en lecture de mots irréguliers, avec des scores plus faibles observés dans les groupes bilingues. Ces résultats pointent le rôle déterminant du milieu sur les compétences langagières, permettant ainsi d'expliquer les écarts persistants entre le niveau de littéracie dans les REP/REP+ et celui des autres écoles.

Le rôle des connaissances morphologiques en lecture

Si de nombreuses études sur la lecture se sont attachées à mettre en évidence les liens entre le profil SES des élèves, leurs compétences langagières et l'acquisition de la lecture, d'autres travaux ont pointé le rôle crucial de la morphologie pour le développement de ces apprentissages. Nous nous concentrons sur la morphologie dérivationnelle, qui correspond au mécanisme de construction des unités lexicales

(*e.g.*, à partir du lexème [LANGUAGE] sont dérivés les mots formes *langage*, *langagier*, *languette*). En effet, à ce jour, une grande majorité d'études en psycholinguistique a montré que les adultes, tout comme les adolescents, sont tous sensibles à la structure morphologique des mots lorsqu'ils traitent la langue écrite (*e.g.*, Dawson *et al.*, 2018).

La conscience morphologique (CM), définie comme l'habileté à identifier et à manipuler des morphèmes, les plus petites unités de sens dans les mots, constitue un facteur essentiel pour l'efficacité en lecture (*e.g.*, Carlisle, 2000 ; Kuo & Anderson, 2006). Cette connaissance métalinguistique est, avec la conscience phonologique¹, l'un des corrélats les plus forts de réussite dans le processus de lecture et de compréhension : Les jeunes lecteurs avec un niveau élevé de connaissances morphologiques analysent le sens des mots morphologiquement complexes plus efficacement, ce qui entraîne des bénéfices en cascade pour l'efficacité de la compréhension de l'écrit (*e.g.*, Deacon & Kirby, 2004 ; Colé *et al.*, 2004 ; Levesque *et al.*, 2022 ; Deacon & Levesque, 2024 ; Kirby *et al.*, 2012, 2025) et cette relation entre les connaissances morphologiques et la compréhension en lecture deviendrait plus forte entre le CM1 et le CM2 (*e.g.*, Fejzo, 2021). Un nombre croissant de preuves empiriques montre enfin que les compétences morphologiques jouent non seulement un rôle chez les enfants monolingues mais aussi chez les bilingues (*e.g.*, Jeon & Yamashita, 2014).

Toutefois l'efficacité en lecture appelle à distinguer le décodage de l'écrit — compétence instrumentale permettant l'accès au lexique *via* l'automatisation des correspondances graphème-phonème, soutenue par la conscience phonologique — des processus de compréhension, qui mobilisent non seulement l'activation des représentations lexicales, mais aussi leur intégration syntaxique au niveau de la phrase, ainsi que leur articulation à l'échelle du texte, comme le décrivent les modèles de la compréhension en lecture (*e.g.*, Stafura & Perfetti, 2017 ; Kim, 2020 ; Levesque *et al.*, 2021). Or, les bons et les mauvais « compreneurs » se distinguent surtout par leurs capacités à extraire des indices à partir de leurs compétences inférentielles (*e.g.*, Levesque *et al.*, 2019). Ainsi lors de la lecture fluide, les apprenants qui ont efficacement acquis des compétences morphologiques accèdent plus rapidement et automatiquement aux représentations mentales des mots et sont donc capables de faire des inférences plus efficaces sur le sens général, améliorant ainsi leur compréhension à la fois au niveau local (lors de la rencontre d'un mot peu ou pas familier) et au niveau général du texte (en en saisissant le sens littéral et implicite).

Dans la littérature, la CM est généralement considérée comme une forme explicite des connaissances morphologiques. Celles-ci peuvent en effet être distinguées selon leur degré d'explicitation en deux grands types : les connais-

1. La sensibilité et l'habileté à manipuler les sons des mots (*e.g.*, BRADLEY & BRYANT, 1983 ; GOSWAMI & BRYANT, 1990 ; de JONG & VAN DER LEIJ, 2002 ; CASTLES & COLTHEART, 2004 ; LANDERL *et al.*, 2019).

sances implicites et les connaissances explicites. Les connaissances implicites sont inconscientes, tacites, qui ne sont pas utilisées délibérément mais activées pendant le traitement du langage. Elles se développent à travers l'exposition à la langue, avec le temps et l'expérience et reposent sur le système de mémoire procédurale, induisant un traitement morphologique spontané. Les connaissances explicites, en revanche, renvoient à ce que nous savons consciemment sur la structure des mots, des connaissances pouvant être utilisées de manière stratégique et développées par un enseignement explicite et un effort conscient. Elles reposent donc sur la mémoire déclarative et concernent la CM. Toutefois, comme souligné par Nagy *et al.* (2013) « la frontière entre le traitement morphologique et la conscience morphologique n'a pas souvent été discutée dans la littérature de recherche sur la conscience morphologique; elle n'est pas non plus nécessairement une frontière facile à établir » (p. 4, notre traduction). Cependant cette distinction est pertinente pour le développement de l'alphabétisation en langue seconde (L2), car la L2 est généralement moins implicitement développée (et moins automatisée) que la langue première (L1), en particulier si la L2 s'acquiert dans le cadre d'un enseignement formel, comme à l'école. Étant donné que la recherche sur l'acquisition de la langue indique que ces deux dimensions de connaissances se soutiennent mutuellement, il est important de comprendre comment une telle coopération entre l'explicite et l'implicite se manifeste-t-elle à travers le développement des compétences morphologiques des apprenants, dont la langue de scolarisation n'est pas toujours la langue maternelle (entre 20 et 40 % des élèves en France sont plurilingues selon Bijeljac-Babic & Auger, 2025).

Corbin énonçait déjà en 1991, ce développement progressif de compétences: « La compétence lexicale se raffine avec l'âge, s'enrichit de connaissances spécifiques nouvelles, de bases nouvelles et d'idiosyncrasies en tout genre, mais pour l'essentiel, la compétence dérivationnelle est construite dès l'enfance » (p. 102) et rajoutait plus loin: « La compétence métalexicale s'exprimerait également de façon spécifique dans le fait que, alors que dans la pratique habituelle de la langue, les règles s'appliquent sans que le locuteur en soit conscient, le locuteur et l'auditeur deviennent conscients de l'application des règles lors de la perception de mots "nouveaux" » (p. 112). Ainsi, l'autrice ne décrit pas seulement le développement de la compétence dérivationnelle dans le cadre plus large de la compétence lexicale, elle propose aussi une réflexion sur l'articulation entre les connaissances implicites et explicites. À travers la notion de compétence métalexicale, elle distingue clairement l'usage inconscient des règles morphologiques et la prise de conscience de ces règles dans certaines situations langagières, comme l'exposition à des mots nouveaux, à l'oral comme à l'écrit.

Diverses dimensions des compétences métalexicales participent au développement le conscience morphologique (*e.g.*, Deacon *et al.*, 2008): Le décodage morphologique (*i.e.*, capacité à utiliser des morphèmes pour lire correctement un mot), la conscience de la structure morphologique (*i.e.*, la capacité à recon-

naître la structuration interne d'un mot), l'analyse de la morpho-sémantique et de la morpho-syntaxe (*i.e.*, la capacité de déduire le sens « probable » d'un mot complexe à partir de celui de ses constituants morphologiques et de déduire sa fonction syntaxique à partir du suffixe, *via* l'analogie avec d'autres mots partageant le même suffixe). Ces compétences morphologiques ont été identifiées comme parmi les corrélats les plus forts de la réussite en lecture. Les jeunes lecteurs possédant des connaissances morphologiques élevées sont mieux équipés pour analyser le sens et la fonction des mots morphologiquement complexes, engendrant des avantages significatifs pour la compréhension de la lecture (*e.g.*, Deacon *et al.*, 2010). De plus, cette relation est réciproque : une exposition plus intensive à l'imprimé conduit à des connaissances morphologiques plus solides, et inversement. La CM étant une construction multidimensionnelle, des travaux de recherche se sont attachés à systématiquement administrer aux apprenants plusieurs tâches différentes en partant de l'hypothèse selon laquelle il existerait des profils différenciés de CM et que chaque sous-composante contribuerait à la littéracie à des niveaux distincts. (*e.g.*, Goodwin *et al.*, 2020 ; Levesque *et al.*, 2021 ; Appel *et al.*, 2022). Leurs résultats ont montré que les élèves présentaient des profils de développement variés de forces et de faiblesses, selon les différentes tâches morphologiques utilisées, suggérant ainsi l'existence de « degrés » de CM au cours du développement langagier. D'autres études ont enfin souligné que les très jeunes lecteurs (*i.e.*, GSM, CP, CE1) ou ceux ayant un développement langagier atypique, comme les dyslexiques, tirent aussi profit de leur CM lors de l'acquisition de la lecture (*e.g.*, Sparks & Metsala, 2023 ; Giazitzidou & Padeliadu, 2022 ; Sparks & Deacon, 2018 ; Deacon *et al.*, 2019, 2008 ; Diamanti *et al.*, 2017 ; Breadmore & Carroll, 2016 ; Burani, 2010 ; Casalis *et al.*, 2004 ; Colé *et al.*, 2004). Parce que le développement de la CM pourrait compenser certaines difficultés lexicales ou syntaxiques, fréquentes chez les élèves bilingues parlant une langue minoritaire à la maison et/ou issus de milieux défavorisés, elle mérite une attention particulière dans ce contexte.

L'objectif de notre contribution est d'examiner le rôle critique de différentes dimensions de la CM pour l'efficacité en lecture chez des élèves de fin CE1 (7-8 ans) scolarisés en REP. Nous nous appuyons sur l'hypothèse principale formulée par la littérature selon laquelle, la CM a un effet sur la compréhension de la lecture. L'objectif de cette première étude « pilote », du fait de la taille restreinte de notre échantillon (N = 10) et du nombre d'épreuves soumises aux participants, vise ainsi à amorcer une meilleure compréhension des relations complexes entre le Statut économique et social (SES), les compétences morphologiques et la compréhension de texte, afin d'éclairer la problématique d'une littéracie moins développée chez les apprenants issus de milieux socio-économiques défavorisés. Dans ce cadre, nous mettons un accent particulier sur la morphologie dérivationnelle du français qui représente un champ d'investigation privilégié pour l'étude des processus de création des mots, à la fois en raison de sa variation morphologique large

et complexe et de sa richesse. Cette langue présente également la particularité d'avoir une orthographe opaque, où les correspondances phonèmes-graphèmes ne sont pas toujours consistantes. Par exemple, le son [o] peut être orthographié de 3 façons : *o*, *au*, *eau* (e.g., Giraudo, 2018). Cette difficulté supplémentaire, par rapport à d'autres langues dont l'orthographe est transparente (e.g., italien, espagnol, allemand), pourrait toutefois être compensée par la CM qui jouerait un rôle d'autant plus important pour les langues dites opaques. En effet, le décodage de l'écrit des langues formellement opaques pourrait s'appuyer sur une stratégie de segmentation morphologique des mots complexes, plutôt que sur des mises en correspondances graphème-phonème.

Dans notre étude, outre l'usage de sept tests standardisés permettant de mesurer les compétences verbales et non-verbales des élèves afin de les situer par rapport à ceux du même âge et niveau scolaire, nous examinons conjointement quatre dimensions de leur connaissances morphologiques (*i.e.*, analyse morpho-sémantique, analyse morpho-syntaxique, connaissances relationnelles et connaissances distributionnelles). Par ailleurs, afin d'éprouver les effets de ces connaissances sur la compréhension de texte, nous mesurons leur compréhension de texte et nous cherchons à savoir quelles connaissances morphologiques sont corrélées avec leur capacité à comprendre le sens littéral et implicite que véhicule un texte narratif. Ces données expérimentales sont complétées par des données déclaratives concernant leur profil sociolangagier, grâce à l'usage d'un questionnaire adressé à l'élève et d'un autre, à ses parents/représentants.

Méthodologie expérimentale de l'étude présentée

Participants. 10 élèves recrutés dans une classe de CE1 d'une école élémentaire située dans le REP de Decazeville (département de l'Aveyron en France), âgés en moyenne de 7,7 ans (écart type 0.3 ans), ont participé à notre étude¹. Aucun ne présentait de troubles langagiers ou de troubles d'apprentissage identifiés. Le Statut Économique et Social (SES) des élèves a été défini comme faible sur la base de leur lieu de résidence, situé dans une zone péri-urbaine sensible, caractérisée par une forte proportion de familles bénéficiant d'aides financières de l'État.

Outils de mesure et matériels linguistiques. Les outils et matériels sont constitués de deux questionnaires, de sept tests standardisés, de quatre épreuves morphologiques et d'une épreuve de compréhension de texte.

Consentement à participation et questionnaires sociolangagiers. En amont de leur participation à notre étude, les responsables légaux des élèves ont donné leur autorisation à participation de leur enfant sur la base d'une notice d'information et de consentement éclairé, qui a été validée par le Comité d'Éthique de la Recherche

1. La taille de l'échantillon limitée à 10 élèves est justifiée par le fait que la typologie de cette école induit des effectifs très réduits d'élèves par classe.

de l'université de Toulouse¹. Afin, de recueillir des informations sur le profil de chaque élève, une fois leur consentement donné, leurs représentants ont rempli un formulaire comportant des questions générales sur l'histoire langagière et les compétences (8 items), le parcours scolaire (5 items), le développement langagier (3 items) de l'élève. Ces données complétées par des questions sur le profil éducatif et langagier des représentants (3 items). Afin d'obtenir une mesure la plus ajustée possible du degré d'exposition et d'usage de la langue de l'élève, un questionnaire oral de 7 questions (*e.g.*, « À la maison, quand tes parents te félicitent ou te grondent, c'est : a. Surtout en français, b. Surtout dans une autre langue, c. Deux langues »), portant sur ses habitudes langagières en français en dehors de l'école, lui a été soumis directement avant le début de la passation des tests standardisés et épreuves morphologiques.

Tests standardisés

Fluidité en lecture. Le test Alouette R (Lefavara, 2005) a permis de mesurer les capacités de décodage et de fluence en lecture. L'élève doit lire à voix haute un texte de 265 mots pendant 3 minutes maximum.

Lecture de mots et pseudo-mots. Une liste de 20 mots réguliers (*e.g.*, *vague*), une de 20 mots irréguliers (*e.g.*, *ville*) et une autre de 20 pseudo-mots (*e.g.*, *givor*) issus de la batterie ODEDYS (Jacquier-Roux *et al.*, 2005) sont présentées à l'écrit à l'élève. Sa tâche consiste à lire les items un par un.

Répétition de mots et pseudo-mots. Une liste de 16 mots (*e.g.*, *spectacle*) et une liste de 20 pseudo-mots (*e.g.*, *koguchi*) issues de la batterie ODEDYS sont présentées oralement à l'élève afin d'évaluer sa mémoire phonologique à court terme. La tâche consiste à répéter les items un à un.

Conscience phonologique. Pour évaluer cette compétence, deux sous-tests de la batterie ODEDYS ont été utilisés :

- Tâche de fusion : Écouter deux mots successifs (*e.g.*, [ʃjẽ] *chien* et [akʁupi] *accroupi*) et utiliser le premier phonème de chaque mot ([ʃ] et [a]) pour les fusionner en un item (*e.g.*, [ʃa]).
- Tâche de suppression : Écouter un mot (*e.g.*, [kan] *cane*) et produire un nouveau mot en supprimant le premier phonème (*e.g.*, [an] *ane*).
- Chaque tâche comporte 10 items.

Mémoire non verbale. Un sous-test de la batterie ODEDYS a été utilisé pour évaluer la mémoire à court terme et la mémoire de travail. L'élève rappelle des séries de chiffres (11 séries croissantes de chiffres, allant de 2 à 8 chiffres) d'abord dans l'ordre puis à l'envers.

Connaissance du vocabulaire. Le vocabulaire en français a été mesuré à l'aide d'une version adaptée du test de vocabulaire en images *Peabody* (EVIP A,

1. Avis n° 2023-632

Exemple de phrase à compléter :	Réponses possibles :	
« Le mois prochain, on organise une ___ à l'école. »	a. tuirade (N)	✓
	b. tuirement (ADV)	×
	c. tuirir (V)	×
	d. tuirible (ADJ)	×

Connaissances relationnelles et distributionnelles. Basée sur les travaux Carlisle (2000), cette tâche mesure la sensibilité à la structure morphologique des mots complexes et aux liens sémantico-formels que les mots de la même famille morphologique entretiennent au travers de deux sous-tâches. 30 phrases à trou commençant par un mot indice sont complétées par l'élève. La moitié de ces phrases porte sur ses connaissances distributionnelles (a), l'autre sur ses connaissances relationnelles (b) :

A. CONNAISSANCES DISTRIBUTIONNELLES	Produire un mot dérivé à partir d'une base
Exemple	Indice: <i>orage</i> . Phrase: « En hiver, le ciel est souvent <u>orageux</u> . »
B. CONNAISSANCES RELATIONNELLES	Retrouver /extraire la base d'un mot dérivé.
Exemple	Indice: <i>chaton</i> . Phrase: « Mon voisin aime beaucoup son <u>chat</u> . »

Compréhension de texte. La compréhension de texte a été évaluée par une tâche de lecture silencieuse d'un texte narratif adapté au niveau scolaire des élèves, suivie d'une série de 26 questions à choix de réponse multiple (3 réponses courtes possibles), la moitié des questions portait sur le sens littéral du texte, l'autre sur son sens implicite. Afin d'évacuer la problématique de la mémoire à court terme, les élèves disposent du texte écrit pour éventuellement s'aider à répondre aux questions.

Procédure de passation. Les élèves ont été évalués individuellement en commençant par le questionnaire oral sur leurs pratiques langagières, ils ont ensuite été soumis aux tests standardisés puis aux épreuves morphologiques. Enfin, la compréhension de texte a été évaluée collectivement en classe. Les passations se sont déroulées en avril-mai 2023, soit au 4^e trimestre de la classe de CE1.

Résultats. Les analyses statistiques inférentielles ont été réalisées avec le logiciel *Jamovi* (2022).

- a Profils langagiers. Comme illustré dans le tableau 1, l'analyse descriptive des réponses aux questionnaires soumis, d'un côté aux représentants des élèves et de l'autre, à l'élève ont mis en évidence que si 40 % d'entre eux possèdent une langue autre que le français, l'usage du français est majoritaire tant du point de vue des parents (88,5 %) que de l'élève (84,8 %). Les données du tableau 1 ont été mises en corrélation avec les données langagières recueillies. Toutefois, l'analyse de corrélation, réalisée avec le test statistique *R* de Bravais-Pearson, n'a révélé aucune corrélation significative (tous les $p > 0,10$).

Tableau 1 — Langues 1 et 2 et usages moyens exprimés en pourcentage (%) selon les élèves et leurs parents

Langues 1 et 2	% d'usage
Langue seconde autre que le français	40 %
Usage du français hors école (parents)	88,5 %
Usage de la langue seconde (parents)	26,9 %
Usage du français hors école (élève)	84,8 %

b Tests standardisés. Les scores des élèves de notre échantillons ont été comparés à ceux des élèves de mêmes âge et niveau scolaire (Cf. tableau 2), la significativité des différences a été estimée à l'aide d'un test statistique *t* de Student appliqué à des mesures non répétées.

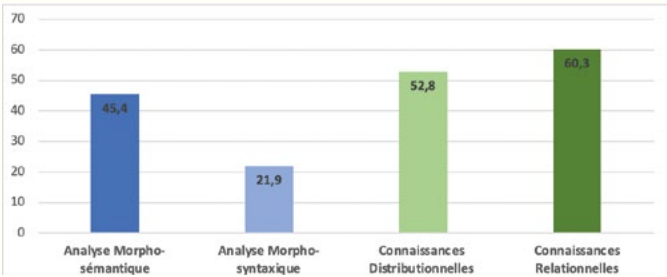
Tableau 2 — Scores moyens des élèves comparés aux scores standards

Épreuves standardisées	Scores élèves	Scores standards	Significativité
Lecture à haute voix	87,6	93	$p < ,02^*$
Lecture de mots réguliers	18,5	18,1	$p > ,10$
Lecture de mots irréguliers	13,3	14,5	$p < ,02^*$
Lecture de pseudo-mots	15,3	15,3	$p > ,05$
Répétition de mots	14,7	15,5	$p < ,05^*$
Répétition de pseudo-mots	14	18,4	$p < ,002^*$
Suppression de phonèmes	5,3	6,2	$p > ,05$
Fusion de phonèmes	8,1	6,5	$p < ,001^*$
Vocabulaire réceptif	835	102,1	$p < ,0007^*$
Compréhension orale sémantico-syntaxique	9	9,5	$p > ,10$
Mémoire à court terme	4,9	4,6	$p > ,05$
Mémoire de travail	3,4	4,5	$p < ,0007^*$

* $p < 0,05$

Les résultats ont révélé que notre échantillon présente des performances plus faibles comparées aux scores standards, pour la fluidité de la lecture d'un texte, la lecture de mots irréguliers, une partie de la conscience phonologique (*i.e.*, fusion de phonèmes) et le vocabulaire réceptif. Nos participants ont également une mémoire de travail (*i.e.*, répétition de séries de chiffres à l'envers) et une mémoire phonologique à court terme (*i.e.*, répétition de mots et pseudo-mots) plus faibles. Toutefois, leur lecture de mots réguliers et pseudo-simples n'est pas différente des autres élèves, de même que leur compréhension orale, leurs compétences sémantico-syntaxique et leur mémoire à court terme non verbale.

c Connaissances morphologiques. L'analyse descriptive des différents scores moyens recueillis dans chacune des quatre dimensions morphologiques mesurées révèlent des pourcentages inférieurs à 60 % de réussite (Cf. Graphique 2).



Graphique 2 – Scores (en %) aux 4 épreuves morphologiques.

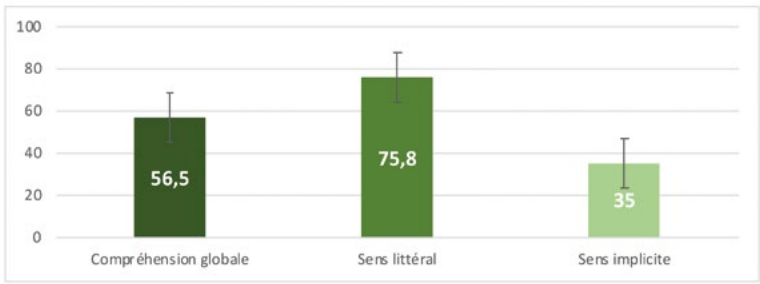
L'analyse statistique des comparaisons spécifiques (*t* de Student) présentée dans le Tableau 3, montre que l'analyse morpho-syntaxique, épreuve la moins bien réussie (21,9%), diffère significativement des 3 autres dimensions mesurées. Les connaissances relationnelles (52,8%) sont moins bien maîtrisées que les connaissances distributionnelles (60,8%).

Tableau 3 – Comparaisons des scores moyens aux 4 épreuves morphologiques

Comparaisons		Valeur <i>t</i>	ddl	<i>p</i>
Morpho-sémantique	Morpho-syntaxe	5,122	9	<,001*
	Connaissances relationnelles	0,301	9	0,770
	Connaissances distributionnelles	1,837	9	0,307
Morpho-syntaxe	Connaissances relationnelles	3,750	9	<,005*
	Connaissances distributionnelles	4,600	9	<,001*
Connaissances distribu- tionnelles	Connaissances relationnelles	2,255	9	0,051**

*différence significative ; ** tendance à la significativité

d Compréhension de texte. L'analyse descriptive des scores moyens obtenus fait apparaître un taux de compréhension globale moyennement satisfaisant (56,5 %) qui se compose d'un taux de compréhension du sens littéral élevé (75,8 %) mais d'un taux de compréhension du sens implicite faible (35 %) comme l'illustre le Graphique 3.



Graphique 3 — Scores moyens (en %) en compréhension de texte.

L'application d'un test statistique (*t* de Student pour échantillons appariés) montre une différence significative entre les deux composantes de la compréhension en lecture ($t(9) = 3,25, p < 0,01$). Saisir l'implicite d'un texte semble constituer pour ces élèves un véritable défi en termes de compréhension de l'écrit.

Les compétences orales en matière de vocabulaire et de compréhension sémantico-syntaxique de l'élève étant liées à la compréhension de texte, nous avons examiné la nature des corrélations entre les compétences langagières issues des tests standardisés (*i.e.*, EVIP et ECOSSE, respectivement) et la compréhension en lecture d'un texte (Cf. tableau 4) :

Tableau 4 — Coefficients de corrélation entre la connaissance du vocabulaire, la compréhension sémantico-syntaxique et les composantes de la compréhension de texte

	Épreuves orales		Épreuve écrite		
	Vocab. (Test EVIP)	Compréhension sém-syntaxique (Test ECOSSE)	Compréhension globale du texte	Sens Littéral	Sens Implicite
Vocabulaire	-				
Compréhension sém-syntaxique	-0,368	-			
Compréhension globale du texte	0,653*	0,726*	-		
Sens Littéral	0,353	0,170	0,519	-	
Sens Implicite	0,597***	0,756*	0,914**	0,129	-

* $p < 0,05$; ** $p < 0,01$; *** $0,05 < p < 0,10$

La compréhension globale du texte corrèle positivement avec les deux compétences orales. Le sens implicite est positivement lié à la compréhension sémantico-

syntactique et partiellement lié à la connaissance du vocabulaire. Ces corrélations suggèrent que meilleures sont les compétences verbales à l'oral, meilleure est, en lecture, la compréhension d'un texte.

- e Liens entre scores aux épreuves standardisées et connaissances morphologiques. Afin de mesurer cette fois, l'impact des connaissances en matière de vocabulaire et de la compréhension sémantico-syntactiques des élèves sur la qualité de leurs connaissances morphologiques, une analyse de corrélation a été réalisée (R de Bravais-Pearson). Des corrélations significatives différenciées sont apparues (cf. tableau 5).

Tableau 5 — Coefficients de corrélation entre les résultats à deux tests standards et les connaissances morphologiques

	Connaissance du vocabulaire (Test EVIP)	Compréhension sémantico- syntactique (Test ECOSSE)
Analyse morpho-sémantique	0,265	0,240
Analyse morpho-syntactique	0,190	0,069
Connaissances relationnelles	0,787**	0,337
Connaissances distributionnelles	0,713*	0,704*

* $p < 0,05$; ** $p < 0,01$; $0,05 < p < 0,10$

Il a été constaté que les connaissances distributionnelles et relationnelles dépendent de la connaissance du vocabulaire des élèves. Ce résultat semble logique, plus l'élève a de vocabulaire plus les familles et séries morphologiques sont riches et structurées. La capacité à dériver à partir d'une base lexicale donnée, le mot suffixé qui complète correctement une phrase du point de vue sémantique et syntactique (*i.e.*, connaissances distributionnelles) est corrélée avec la compréhension sémantico-syntactique des élèves. De nouveau, ce résultat nous semble cohérent, les élèves qui repèrent correctement les situations imaginées que suggèrent des propositions plus ou moins complexes (*i.e.*, Test ECOSSE), sont naturellement ceux qui sont le plus à l'aise pour compléter correctement une phrase à l'aide d'un mot suffixé.

Liens entre connaissances morphologiques et compréhension de texte. Afin d'éprouver l'hypothèse de la littérature selon laquelle les connaissances morphologiques ont des effets bénéfiques sur la compréhension en lecture, nous avons réalisé une analyse des corrélations (R de Bravais-Pearson) entre les taux de réussite aux quatre épreuves morphologiques et ceux en compréhension globale, littérale et implicite du texte proposé aux élèves.

Tableau 6 — Coefficients de corrélation entre connaissances morphologiques et compréhension de texte

	Morpho- sém.	Morpho- synt.	Relatio.	Distrib.	Comp. Globale	Sens Litt.	Sens Impl.
Analyse morpho-sémantique	-						
Analyse morpho-syntaxique	0,261	-					
Connaiss. Relationnelles	0,307	0,200	-				
Connaiss. Distributionnelles	0,032	-0,256	0,756*	-			
Compréhension Globale	0,130	0,149	0,291	0,509***	-		
Sens Littéral	0,097	0,337	0,162	-0,035	0,519	-	
Sens Implicite	0,103	0,016	0,261	0,608*	0,914**	0,129	-

*p < 0,01 ; ** p < 0,001 ; *** 0,05 < p < 0,10

Nous observons une corrélation positive qui tend à être significative entre les connaissances distributionnelles et la compréhension globale d’un texte et qui est significative avec son sens implicite. De bonnes connaissances sur les relations entre les bases lexicales et ses dérivés favorisent une bonne compréhension du sens implicite d’un texte, engendrant ainsi des bénéfices sur sa compréhension globale. La compréhension globale étant corrélée positivement avec celle de son sens implicite, ce résultat suggère également que les élèves les plus capables d’accéder au sens implicite, et donc de mobiliser une lecture plus profonde et interprétative, sont aussi ceux qui obtiennent les meilleurs scores en compréhension globale. Ces résultats confortent l’idée selon laquelle la capacité à traiter l’implicite constitue un levier important pour une lecture efficace, et qu’elle peut être soutenue par des connaissances morphologiques bien développées — en particulier dans leur dimension distributionnelle.

Au cours de l’apprentissage de la lecture, des liens bilatéraux se développent entre les bases et leurs dérivés morphologiques, qui se structurent en familles morphologiques dans le lexique mental des élèves, comme le montre la corrélation positive entre les connaissances distributionnelles et relationnelles.

Discussion

Dans cet article consacré à la problématique du recul de la littéracie en France, nous nous sommes intéressées aux facteurs que la littérature scientifique a identi-

fiés comme potentiellement prédicteurs de la réussite en lecture et compréhension. Dans ce cadre, nous avons centré notre étude sur un échantillon d'élèves de CE1 scolarisés en REP afin de sélectionner des apprenants au profil sociolangagier vulnérable et qui sont potentiellement les premiers touchés sur le plan de leurs compétences en français. Leurs difficultés les plus marquées sont à la fois soulignées par les différents rapports de l'OCDE et du MEN mais aussi par les travaux de littérature qui ont montré un impact direct du Statut Économique et Social (SES) sur les performances des élèves en lecture. Par ailleurs en dehors de cette problématique, des travaux expérimentaux sur les processus d'acquisition de la lecture ont pu démontrer le rôle crucial de la conscience morphologique pour sa réussite, laquelle se développe tout au long de son apprentissage. Toutefois, ces travaux ont souligné que les connaissances morphologiques impliquaient différentes dimensions ayant des effets spécifiques à chacun des niveaux que la lecture implique et que ces effets se manifesteraient de manière différenciée au cours du développement langagier des élèves du CP au CM2 (*i.e.*, de 6 à 11 ans). Ce dernier aspect souligne donc l'importance de mesurer conjointement plusieurs dimensions des connaissances morphologiques pour chaque niveau scolaire donné. Des élèves de fin de CE1 (7,7 ans en moyenne), pour lesquels les processus décodage de l'écrit — débutant généralement en France vers l'âge de 5 ans — ont déjà été automatisés, ont été testés individuellement sur leurs capacités à manipuler la structure morphologique des unités complexes, à en extraire du sens, à en comprendre la fonction grammaticale et à établir des liens sémantico-formels entre les mots apparentés sur le plan morphologique. Enfin, ils ont été testés sur leur compréhension de texte. Des questionnaires adressés à leurs parents ou directement à eux, sur leur degré d'exposition au français, langue de scolarisation, et à une éventuelle autre langue parlée en dehors de l'école, nous ont permis d'établir leur profil langagier.

Trois hypothèses ont servi de cadre méthodologique à nos travaux : Le statut économique et social affecte les compétences langagières et extra-langagières de ces élèves. De ce fait, cette première hypothèse a pour corollaire une seconde hypothèse selon laquelle leur conscience morphologique est amoindrie. D'où découle une troisième hypothèse selon laquelle leur compréhension de la lecture en est directement touchée. En effet, une moins bonne CM ralentit l'accès au lexique, ce qui a pour conséquence de davantage consacrer leur charge cognitive au décodage du texte plutôt qu'à sa compréhension.

Les données que nous avons recueillies à l'aide de tests standardisés indiquent que ces élèves aux profils fragiles ont en effet des compétences générales plus faibles que celles de leurs pairs (*i.e.*, scores standardisés), ce malgré une large exposition et pratique du français dans leur environnement quotidien (scolaire et extra-scolaire > 80 %), comme les données des questionnaires l'indiquent. Ces résultats sont cohérents avec les études antérieures ayant montré que les enfants issus de milieux socio-économiques défavorisés tendent à obtenir des performances plus faibles aux tâches de raisonnement non verbal, ce qui pourrait refléter des diffé-

rences d'exposition à certaines formes de stimulation cognitive (*e.g.*, Bellocchi & Bonifacci, 2024). Nos données rajoutent que nos élèves accusent également un recul en matière de conscience phonologique (*i.e.*, fusion de phonèmes), lecture à haute voix d'un texte et lecture de mots irréguliers. Il semble dès lors que la nature du langage auquel ces élèves sont exposés — marqué par une faible diversité et un accès restreint à une langue soutenue et complexe, notamment à l'écrit — soit un déterminant majeur de leurs compétences en lecture.

Les performances des élèves aux différentes épreuves morphologiques varient selon la sous-compétence ciblée, ce qui confirme que la conscience morphologique est une construction multidimensionnelle composée de sous-habiletés distinctes, contribuant différemment à la compréhension en lecture. Plus précisément, les habiletés relationnelles (extraire la base d'un mot suffixé) et distributionnelles (produire un mot suffixé à partir d'une base) semblent être acquises plus précocement et mobilisées plus facilement que les habiletés qui requièrent un raisonnement métalinguistique ou une sensibilité aux dimensions sémantiques et syntaxiques du langage. Cela se reflète dans les taux de réussite : 60 % pour les connaissances relationnelles et 52 % pour les distributionnelles, contre seulement 45 % pour l'analyse morpho-sémantique et 22 % pour l'analyse morpho-syntaxique. L'ensemble de ces résultats suggère un degré de conscience morphologique, situé autour de 45,7 %, et révèle un lexique mental qui n'est pas encore assez structuré en familles et séries morphologiques. Enfin, les scores obtenus aux épreuves morphologiques sont positivement corrélés, bien que de manière différenciée, aux résultats des tests standardisés mesurant les compétences langagières générales à l'oral (vocabulaire, mémoire à court terme et compréhension sémantico-syntaxique), confirmant l'interdépendance des habiletés cognitives et linguistiques impliquées dans les processus de lecture et de compréhension de texte.

Dans notre étude, la seule corrélation significative observée entre les performances en lecture et les compétences morphologiques concerne la relation entre les connaissances distributionnelles et la compréhension du texte, en particulier sa dimension implicite. Le fait que les autres dimensions de la conscience morphologique ne montrent pas de lien statistiquement significatif avec la compréhension pourrait s'expliquer par leur faible niveau de développement chez les élèves de notre échantillon. Ainsi, nos résultats confirment partiellement les liens établis dans la littérature entre conscience morphologique et compréhension en lecture, tout en laissant penser que, dans le cas d'élèves aux profils langagiers fragiles, la faiblesse conjointe des compétences morphologiques et de lecture pourrait limiter l'émergence de relations plus étendues entre ces deux domaines.

Ainsi, dans notre échantillon, les compétences morphologiques ne soutiennent que partiellement la compréhension en lecture, tout comme les compétences langagières générales, elles aussi fragiles. Le statut économique et social semble impacter négativement certaines compétences orales, tant langagières (comme le vocabulaire ou la compréhension des structures grammaticales) qu'extra-

langagières. Dans ce contexte, les connaissances morphologiques des élèves, encore insuffisamment développées, ne parviennent pas à compenser ces fragilités, limitant ainsi leur contribution à l'élaboration du sens lors de la lecture.

Du point de vue de la théorie de la charge cognitive, concept multidimensionnel se référant à la quantité limitée de ressources cognitives mobilisées en mémoire pour réaliser une tâche d'apprentissage (*e.g.*, Sweller *et al.*, 1998), nous émettons l'hypothèse que la majorité des ressources cognitives de ces élèves, du fait de leur profil sociolangagier et de leur mémoire de travail non verbale moins développée que les autres élèves, a été consacrée au décodage et à la compréhension de la surface du texte, saturant ainsi la quantité de ressources mobilisables en mémoire de travail pour réaliser la tâche. La conséquence en est que trop peu de ressources nécessaires aux processus inférentiels liés à la construction-intégration du sens du texte ont pu être mobilisées, expliquant ainsi le taux de compréhension de ces élèves.

Les données des évaluations des acquis des élèves CE2 menées par le MEN en 2024¹ indiquent qu'en début de CE2 (ce qui correspond plus ou moins au niveau scolaire de nos élèves qui ont été testés en fin de CE1), une grande majorité des élèves présente une maîtrise « satisfaisante » pour la compréhension en lecture avec les items « comprendre des phrases lues seul(e) » réussis par 82,5 % d'entre eux et « comprendre un texte lu seul(e) » réussis par 64 % d'entre eux. Dans les réseaux REP, ces taux de réussite passent respectivement à 67,5 % (-15 points) et 50,2 % (-13,8 points). Lorsqu'on met nos données en miroir avec celles du MEN, on peut avancer que notre échantillon semble représentatif de la population des élèves de REP. Leur degré de littéracie, avec une compréhension globale d'un texte lu estimée à 56,5 %, est relativement proche de celle de leurs homologues (50,2 %). Toutefois, les différences entre les matériels, les composantes de la compréhension mesurées (où ne sont pas distingués par le MEN les sens littéral et implicite) et les procédures utilisées ne nous permettent pas de mener des analyses supplémentaires. Il serait donc intéressant de mener une étude expérimentale avec un échantillon plus large afin de mesurer de manière plus précise l'impact de la conscience morphologique de ces élèves sur leur réussite aux évaluations nationales mais également de mesurer le développement des trajectoires des habiletés en lecture au travers d'études longitudinales de cohortes d'élèves (Kirby *et al.*, 2025).

En conclusion, nos résultats corroborent celles énoncées dans la littérature sur les rôles du statut économique et social et de la conscience morphologique pour la lecture, laquelle précise que ces deux facteurs impactent conjointement les traitements sémantique et syntaxique de l'écrit et les processus de construction-intégration du sens d'un texte. Notre étude s'inscrit dans la ligne des travaux de Piccinin et Dal Maso (2023), menés sur l'italien auprès d'un effectif plus important d'élèves

1. <https://www.education.gouv.fr/l-evaluation-des-acquis-des-eleves-en-ce2-des-reperes-de-debut-d-annee-414303>.

de seconde génération de migrants, qui ont montré des liens robustes entre les connaissances morphologiques et les habilités en compréhension de la lecture. Ces résultats mettent en lumière les difficultés des élèves à profil fragile pour l'acquisition de la littéracie et pour lesquels la conscience morphologique doit être développée et renforcée. Elle est en effet garante de la construction d'un lexique mental riche et structuré en familles et séries morphologiques. Nous avançons donc que la morphologie est un facteur d'augmentation du stock des unités lexicales et aide les lecteurs à lire plus rapidement et efficacement les mots familiers et les mots non-familiers sur la base de stratégies métacognitives d'analogies sémantico-formelles entre les mots codés en mémoire à long terme et ceux rencontrés dans les textes. De plus, l'accès rapide au sens et à la fonction grammaticale des mots complexes, qui représentent 80% des mots rencontrés dans les textes en français (Rey-Debove, 1984), a des effets en cascade sur les processus cognitifs d'inférences nécessaires à la construction-intégration d'un modèle mental de la compréhension de texte. Sur le versant de la didactique, les données de la littérature et les nôtres encouragent la mise en œuvre d'activités permettant de soutenir et de développer les différentes dimensions de la conscience morphologique chez tous les élèves mais surtout chez les plus fragiles sur le plan sociolinguistique (*e.g.*, Piccinin & Dal Maso, 2024 ; Berthiaume *et al.*, 2018 ; Fejzo *et al.*, 2016).

Références bibliographiques

- APPEL, Kenn, Victoria S. HENBEST & Yaacov PETSCHER, « Morphological Awareness Performance Profiles of First- Through Sixth-Grade Students », *Journal of Speech, Language and Hearing Research*, 65(3) (2022) : p.1070-1086.
- BELLOCCHI, Stéphanie & Paola BONIFACCI, « Disentangling the impact of bilingualism and SES in literacy skills of language-minority bilingual children and monolingual peers exposed to French », *Reading Psychology*, 45(3) (2024) : p. 219-241.
- BERTHIAUME, Rachel, Daniel DAIGLE & Alain DESROCHERS, *Morphological processing and literacy development: Current issues and research*, Routledge, 2018.
- BIJELJAC-BABIC, Ranka & Nathalie AUGER, *Avoir deux langues et plus à l'école maternelle et primaire. Synthèse de la recherche et recommandations*, Conseil scientifique du ministère de l'Éducation nationale (CSEN), 2025.
- BONIFACCI, Paola, Ida C. FERRARA, Jessica PEDRINAZZI, Francesco TERRACINA & Paola PALLADINO, « Literacy acquisition trajectories in bilingual language minority children and monolingual peers with similar or different SES: A three-year longitudinal study », *Brain Sciences*, 12(5) (2022) : 563 p.

- BONIFACCI, Paola, Giuseppina LOMBARDO, Jessica PEDRINAZZI, Francesco TERRACINA & Paola PALLADINO, « Literacy skills in bilinguals and monolinguals with different SES », *Reading & Writing Quarterly*, 36(3) (2020) : p. 243-259.
- BRADLEY, Lucy & Peter E. BRYANT, « Categorizing sounds and learning to read — a causal connection », *Nature*, 301 (1983) : p. 419-421.
- BREADMORE, Helen L. & Julia M. CARROLL, « Morphological spelling in spite of phonological deficits: Evidence from children with dyslexia and otitis media », *Applied Psycholinguistics*, 37(6) (2016) : p. 1439-1460.
- BURANI, Christina, « Word morphology enhances reading fluency in children with developmental dyslexia », *Linguee Linguaggio*, 9(2) (2010) : p. 177-198.
- CARLISLE, Joanne F., « Awareness of the structure and meaning of morphologically complex words: Impact on reading », *Reading and Writing*, 12(34) (2000) : p. 169-190.
- CALVO, Alejandra & Ellen BIALYSTOK, « Independent effects of bilingualism and socioeconomic status on language ability and executive functioning », *Cognition*, 130(3) (2014) : p. 278-288.
- CASALIS, Séverine, Pascale COLÉ & Delphine SOPO, « Morphological awareness in developmental dyslexia », *Annals of Dyslexia*, 54 (2004) : p. 114-138.
- CASALIS, Séverine, Hélène DEACON & Sébastien PACTON, « How specific is the connection between morphological awareness and spelling in the early and later stages of learning to spell? », *Reading and Writing*, 24 (2011) : p. 491-506.
- CASTLES, Anne & Max COLTHEART, « Is there a causal link from phonological awareness to success in learning to read? », *Cognition*, 91(1) (2004) : p. 77-111.
- CASTLES, Anne, Kathleen RASTLE & Kate NATION, « Ending the reading wars: Reading acquisition from novice to expert », *Psychological Science in the Public Interest*, 19(1) (2018) : p. 5-51.
- COLÉ, Pascale, Carine ROYER, Chrystel LEUWERS & Séverine CASALIS, « Les connaissances morphologiques dérivationnelles et l'apprentissage de la lecture chez l'apprenti-lecteur français du CP au CE2 », *L'année psychologique*, 104(4) (2004) : p. 701-750.
- CORBIN, Danielle, *Morphologie dérivationnelle et structuration du lexique (tomes I et II)*, Presses universitaires du Septentrion, 1991.
- CORBIN, Danielle, « La représentation d'une famille de mots dans le Dictionnaire dérivationnel du français et ses corrélats théoriques, méthodologiques et descriptifs: Autour de la morphologie », *Recherches linguistiques de Vincennes*, 26 (1997) : p. 5-38.

- DAWSON, Nicola, Kathleen RASTLE & Jessie RICKETTS, « Morphological effects in visual word recognition: Children, adolescents, and adults », *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4) (2018): p. 645-654.
- DEACON, Hélène, Emily CAMPBELL, Meredith TAMMINGA & John R. KIRBY, « Seeing the harm in harmed and harmful: Morphological processing by children in Grades 4, 6, and 8 », *Applied Psycholinguistics*, 31(4) (2010): p. 759-775.
- DEACON, Hélène & John R. KIRBY, « Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development », *Applied Psycholinguistics*, 25(2) (2004): p. 223-238.
- DEACON, Hélène & Kyle LEVESQUE, « Mechanisms in the relation between morphological awareness and the development of reading comprehension », *Journal of Educational Psychology*, 116(6) (2024): p. 1052-1069.
- DEACON, Hélène, Rauno PARRILA & John R. KIRBY, « A review of the evidence on morphological processing in dyslexics and poor readers: A strength or weakness? », dans Gavin REID, Angela J. FAWCETT, Franck MANIS & Linda S. SIEGEL (éd.), *The Sage Handbook of Dyslexia*, Sage Publications, 2008, p. 212-237.
- DEACON, Hélène, Shelley Xiuli TONG & Catherine MIMEAU, « Morphological and semantic processing in developmental dyslexia », dans Ludo VERHOEVEN, Charles PERFETTI & Kenneth PUGH (éd.), *Developmental dyslexia across languages and writing systems*, Cambridge University Press, 2019, p. 327-349.
- DE JONG Peter F. & Aryan VAN DER LEIJ, « Effects of phonological abilities and linguistic comprehension on the development of reading », *Scientific Studies of Reading*, 6(1) (2002): p. 51-77.
- DIAMANTI, Vassiliki, Nata GOULANDRIS, Ruth CAMPBELL, R. & Athanassios PROTOPAPAS, « Dyslexia profiles across orthographies differing in transparency: An evaluation of theoretical predictions contrasting English and Greek », *Scientific Studies of Reading*, 22(1) (2017): p. 1-15.
- DIAMANTI, Vassiliki, Angeliki MOUZAKI, Asimia M. RALLI, Faye ANTONIOU, Sofia PAPAIOANNOU, S. & Athanassios PROTOPAPAS, « Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek », *Frontiers in Psychology*, 8 (2017): p. 20-39.
- DIAS-BROENS, Ake S. & Roel VAN STEENSEL, « Home-Visiting in a Shared Reading Intervention: Effects on Children from Low SES and Ethnic Minority Families », *Early Education and Development*, 34(8) (2022): p. 1919-1940.

- DOLEAN, Dacian & Alexandra CĂLUGĂR, « How reliably can we measure a child's true IQ? Socio-economic status can explain most of the inter-ethnic differences in general non-verbal abilities », *Frontiers in Psychology*, 11 (2020) : 2000.
- ELLIS, Nick C., « Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate », dans B. ROBINSON & Nick C. ELLIS (éd.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, Routledge, 2008, p. 372-405.
- FEJZO, Anila, « The contribution of morphological awareness to the spelling of morphemes and morphologically complex words in French », *Reading and Writing*, 29(2) (2016) : p. 207-228.
- FEJZO, Anila, « The contribution of morphological awareness to vocabulary among L1 and L2 French-speaking 4th-graders », *Reading and Writing*, 34 (2021) : p. 659-679.
- FEJZO, Anila, Alain DESROCHERS & Hélène DEACON, « The acquisition of derivational morphology in children », dans Rachel BERTHIAUME, Daniel DAIGLE & Alain DESROCHERS (dir.), *Morphological processing and literacy development: Current issues and research*, Routledge, 2018, p. 112-132.
- FEJZO, Anila, Lucie GODARD & Line LAPLANTE, « Les effets d'une intervention en conscience morphologique sur l'orthographe lexicale chez des élèves arabophones: élèves scolarisés en français dans le 2^e cycle du primaire en milieu défavorisé », *La nouvelle revue de l'adaptation et de la scolarisation*, 76(4) (2016) : p. 45-64.
- GAIZITZIDOU, Sophia & Susana PADELIADU, « Contribution of morphological awareness to reading fluency of children with and without dyslexia: Evidence from a transparent orthography », *Annals of Dyslexia*, 72(3) (2022) : p. 509-531.
- GIRAUDO, Hélène, « Much ado about morphemes », dans Olivier BONAMI, Gilles BOYÉ, Georgette DAL, Hélène GIRAUDO & Fiametta NAMER (éd.), *The lexeme in descriptive and theoretical morphology*, Berlin : Language Science Press, 2018, p. 401-421.
- GOODWIN, Amanda, Yaacov PETSCHER & Jamie TOCK, « Multidimensional morphological assessment for middle school students », *Journal of Research in Reading*, 44(1) (2020) : p. 70-89.
- GOSWAMI, Usha & Peter BRYANT, *Phonological Skills and Learning to Read* (1^{re} éd.), Routledge, 2016.
- JACQUIER-ROUX, Monique, Sylvaine VALDOIS, Michel ZORMAN, Christine LEQUETTE & Guillemette POUGET, *ODEDYS 2: Outil de Dépistage des Dyslexies* (version 2), Grenoble: Laboratoire Cogni Sciences, IUFM de Grenoble, 2005.

- JEON, Eun Hee & Junko YAMASHITA, « L2 reading comprehension and its correlates: A meta-analysis », *Language Learning*, 64(1) (2014) : p. 160-212.
- KIM, Young-Suk Grace, « Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER) », *Journal of Educational Psychology*, 112(4) (2020) : p. 667-684.
- KIRBY, John R., Hélène DEACON, Peter N. BOWERS, Leah IZENBERG, Leslie WADE-WOOLLEY & Rauno PARRILA, « Children's morphological awareness and reading ability », *Reading and Writing*, 25(2) (2012) : p. 389-410.
- KIRBY, John R., Hélène DEACON, S. H., George GEORGIU, Kelly GEIER, Jessica CHAN & Rauno PARRILA, « Effects of morphological awareness, naming speed, and phonological awareness on reading skills from Grade 3 to Grade 5 », *Journal of Experimental Child Psychology*, 253 (2025) : p. 106-188.
- KUO, Li-jen & Richard C. ANDERSON, « Morphological awareness and learning to read: A cross-language perspective », *Educational Psychologist*, 41(3) (2006) : p. 161-180.
- LANDERL, Karine, H. Harald FREUDENTHALER, Moritz HEENE, Peter F. DE JONG, Alain DESROCHERS, George MANOLITSIS, Rauno PARRILA & George K. GEORGIU, « Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency », *Scientific Studies of Reading*, 23(3) (2019) : p. 220-234.
- LECOQ, Pierre, *E.CO.S.SE – Évaluation de la compréhension syntaxico-sémantique*, Paris: Éditions du Centre de Psychologie Appliquée (ECPA), 2011.
- LEFAVRAIS, Pierre, *Test de l'Alouette-R: Manuel pour l'évaluation de la dyslexie*, 2^e éd., Paris: Éditions du Centre de Psychologie Appliquée (ECPA), 2005.
- LEVESQUE, Kyle. C., Helen L. BREADMORE & Hélène DEACON, « How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development », *Journal of Research in Reading*, 44(1) (2021) : p. 10-26.
- LEVESQUE, Kyle C. & Hélène DEACON, « Clarifying links to literacy: How does morphological awareness support children's word reading development? », *Applied Psycholinguistics*, 43(4) (2022) : p. 921-943.
- LEVESQUE, Kyle C., Michael J. KIEFFER & Hélène DEACON, « Inferring meaning from meaningful parts: The contributions of morphological skills to the development of children's reading comprehension », *Reading Research Quarterly*, 54(1) (2019) : p. 63-80.

- NAGY, William E. & Richard C. ANDERSON, « How many words are there in printed school English? », *Reading Research Quarterly*, 19(3) (1984) : p. 304-330.
- NAGY, William E., Joanne F. CARLISLE & Amanda GOODWIN, « Morphological knowledge and literacy acquisition », *Journal of Learning Disabilities*, 47(1) (2013) : p. 3-12.
- OCDE, *Les adultes possèdent-ils les compétences nécessaires pour s'épanouir dans un monde en mutation? : Évaluation des compétences des adultes 2023*, Études de l'OCDE sur les compétences, Paris : Éditions OCDE, 2024.
- PICCININ, Sabrina & Serena DAL MASO, « Investigating factors affecting reading comprehension in second-generation pupils: the role of morphological skills », *Lingue e Linguaggio*, 1 (2023) : p. 117-142.
- PICCININ, Sabrina & Serena DAL MASO, « Alla scoperta delle parole: un intervento didattico per lo sviluppo della consapevolezza morfologica », *Lend, Lingua Nuova Didattica*, 3 (2024) : p. 8-16.
- SPARKS, Erin & Hélène DEACON, « Morphological awareness and vocabulary acquisition: a longitudinal examination of their relationship in English-speaking children », *Applied Psycholinguistics*, 36 (2015) : p. 299-321.
- SPARKS, Erin & Jamie METSALA, « Morphological awareness predicts reading comprehension in first grade students », *Frontiers in Education*, 8 (2023) : p. 1304688.
- STAFURA, Joe & Charles PERFETTI, « Integrating word processing with text comprehension: Theoretical frameworks and empirical examples », dans Kate CAIN, Donald L. COMPTON & Rauno PARRILA (éd.), *Theories of reading development*, John Benjamins, 2017, p. 9-31.
- TAMINE-GARDES, Joëlle, « Introduction à la morphologie (suite) : La morphologie dérivationnelle », *L'Information Grammaticale*, 14 (1982) : p. 31-35.

Associating Form and Meaning in Morphological Theory: Experimental Data From Modern Greek

Nikolaos NTAGKAS

Aristotle University of Thessaloniki

Introduction

The semantics of morphologically complex words constitutes an interesting though rather under-studied area in the field of theoretical morphology (Bauer *et al.*, 2015; Kotowski & Plag 2023; Lieber 2004; Rainer *et al.*, 2014). To a great extent, the relationship between form and meaning in morphological theory remains controversial, having given rise to—and still feeding—the debate between associative *vs.* dissociative approaches.

Within the framework of Associative Construction Morphology (Corbin 1987/1991; 1989; 1991; forthcoming; for Modern Greek see Anastassiadis-Symeonidis 1992; 1998; 2016; 2023; forthcoming; Efthymiou 1999/2012), it is postulated that morphological relationships presuppose not only formal but also semantic relationships, since lexeme construction rules construct concurrently and interdependently both the morphological structure and the predictable meaning of constructed words. Actually, morphological analysis is semantically constrained, as long as morphemes are identified first on the grounds of their meaning (or function) and second on the grounds of their form. Therefore, semantically transparent complex words, *i.e.*, complex words bearing strong semantic relationships with their constituent morphemes, are considered as constructed words with internal morphological structure resulting from the operation of lexeme construction rules. For example, the Modern Greek verb ζυμώνω /zi'mono/ 'to knead', comprising the base morpheme ζυμ-, the derivational verb-forming suffix -ων-, and the inflectional ending -ω, is constructed from the noun base ζύμη /'zimi/ 'dough', comprising the base morpheme ζυμ- and the inflectional ending -η, by means of a lexeme construction rule forming denominal verbs. The same applies in the

case of semantically semitransparent complex words, *i.e.*, complex words bearing mediocre semantic relationships with their constituent morphemes. These words are also treated as constructed, provided that their meanings can be connected to the meanings of their constituent morphemes through semantic processes (*e.g.*, metaphor, metonymy, synecdoche, catachresis, etc.) from the synchronic point of view. For example, the Modern Greek verb ψαρώνω /psa'rono/ 'to psych out', comprising the base morpheme ψαρ-, the derivational verb-forming suffix -ων-, and the inflectional ending -ω, is constructed from the noun base ψάρι /psari/ 'fish', comprising the base morpheme ψαρ- and the inflectional ending -ι, by means of a lexeme construction rule forming denominal verbs and a metaphor changing the literal meaning of the base morpheme ('fish' → 'naive person').

On the other hand, semantically opaque complex words, *i.e.*, complex words bearing negligible semantic relationships with their constituent morphemes, are considered as non-constructed. That is because the semantic relationship between the complex word and its constituents is not synchronically functional, thus pertaining solely to etymology and not to morphology. For example, the Modern Greek verb σκοτώνω /sko'tono/ 'to kill' cannot be constructed from its etymological noun base σκότος /'skotos/ 'darkness', given that we cannot establish a synchronic semantic relationship between them. The etymological noun base σκότος contributes nothing—or, at best, little—to the meaning of the complex word σκοτώνω, and the meaning of the etymological noun base (*i.e.*, 'darkness') cannot be recognized or restored in the meaning of the complex word (*i.e.*, 'to kill'). So, from the associative construction perspective, σκοτώνω is characterized as a non-constructed complex word comprising the derivational suffix -ων-, the inflectional ending -ω, and a base segment σκοτ- that is not a true morpheme, since it lacks meaning, and is not identified as the base morpheme σκοτ- of the noun σκότος. In a similar manner we can treat complex words sharing an incidental pseudo-morphological relationship and not a real etymological relationship with their putative constituents. For example, the Modern Greek verb μαγκώνω /ma'gono/ 'to grip' (originating from an unattested form *μαγκανώνω /maga'nono/ via haplology; see *Dictionary of Standard Modern Greek* 1998) comprises the derivational suffix -ων-, the inflectional ending -ω, and a base segment μαγκ- that bears no meaning and definitely cannot be identified as the homonym base morpheme μαγκ- of the noun μάγκας /'magas/ 'buster'.¹ Hence, semantically opaque complex words, either with or without an etymological base word, are synchronically classified as non-constructed complex words having no genuine base morpheme but a meaningless base segment.

1. As opposed to the semantically transparent complex verb μαγκεύω /ma'gevo/ 'to become/ behave like a buster' (comprising the base morpheme μαγκ-, the derivational verb-forming suffix -εύ-, and the inflectional ending -ω) that is indeed constructed from the noun base μάγκας.

In contrast, according to dissociative theoretical accounts (*e.g.*, Generative Grammar, Aronoff 1976; Chomsky 1970; Chomsky & Halle 1968; Jackendoff 1975; Word Syntax, Di Sciullo & Williams 1987; Lieber 1981; 1992; Selkirk 1982; Toman 1987; 1998; Distributed Morphology, Acquaviva 2009; Halle & Marantz 1993; for Modern Greek see Panagiotidis 2020; 2021; 2023; Panagiotidis *et al.*, 2017), it is assumed that the morphological structure of a complex word is formed first and independent of its meaning, which is after attributed to the complex word through semantic/interpretational rules. This dissociation suggests a form-based morphological analysis undermining the classic definition of morpheme as ‘the smallest individually meaningful element’ (Hockett 1958, 123). Consequently, all morphologically complex words are treated as morphologically structured irrespective of semantic transparency. For example, the above-mentioned Modern Greek complex verbs ζυμώνω /zi'mono/ ‘to kneed’, ψαρώνω /psa'rono/ ‘to psych out’, σκοτώνω /sko'tono/ ‘to kill’, μαγκώνω /ma'gono/ ‘to grip’ are all classified as derived words with the same morphological structure (base morpheme + derivational suffix + inflectional ending) no matter how the meaning of the complex word is related to the meaning of the base morpheme. Namely, from the dissociative perspective, the distinction between constructed and non-constructed complex words is void on the grounds of surface morphological complexity.

Obviously, the semantic transparency of morphologically complex words plays a determinant role within the associative construction framework, inasmuch as the establishment of a semantic relationship between a complex word and its constituent morphemes is a necessary condition for acknowledging a constructed word and its internal morphological structure. Within the dissociative approach though, the role of semantic transparency is understated, as the recognition of internal morphological structure rests mostly in the surface morphological complexity of a word disregarding the relationship of its meaning with the meanings of its constituent morphemes.

The Present Study

In this study we seek to experimentally test the associative vs. dissociative account of derivational morphology in Modern Greek, with a special focus on denominal suffixed verbs and adjectives (see Anastassiadis-Symeonidis 1992; 1998; 2016; Efthymiou 2018; 2023; Efthymiou *et al.*, 2012; Ralli 2022). Our research question is whether meaning affects the mental representation of derived words. A ‘yes’ answer to this question would support the associative account, while a ‘no’ answer would be in favor of the dissociative account.

Accordingly, we designed a reaction time (RT) lexical decision task combined with morphological priming, comprising three critical semantic transparency conditions, *i.e.*, transparent, semitransparent, and opaque (Gonnerman *et al.*,

2007; Hill & Gonnerman 2022; Jared *et al.*, 2017; Kiehl & Joanisse 2011; Marslen-Wilson *et al.*, 2008). Comparing the priming effect (relative to an unrelated control condition) between transparent, semitransparent, and opaque prime-target pairs allows figuring out the effect of semantic transparency and, therefore, answering the research question. In particular, statistically equivalent priming among the transparent, semitransparent, and opaque condition would indicate a purely formal/morphological effect irrespective of meaning, following the dissociative account. Conversely, statistically graded priming among the transparent, semitransparent, and opaque condition would indicate a purely semantic effect irrespective of morphological structure, also consistent with the dissociative account. Instead, statistically equivalent priming between the transparent and semitransparent condition together with null or negative (*i.e.*, inhibitory) priming in the opaque condition would reveal a morphosemantic effect for constructed words in contrast to non-constructed words, indicating that morphological structure presupposes semantic relatedness, as suggested by the associative account.

The task was conducted with adult native speakers of Modern Greek, employing the overt morphological priming paradigm (unmasked visual priming in Experiment 1 and immediate cross-modal priming in Experiment 2), considered to tap into semantic transparency effects at the later stages of morphological processing and in the central levels of lexical representation (Amenta & Crepaldi 2012; Diependaele *et al.*, 2012; Feldman & Milin 2018; Feldman & Moscoso del Prado Martín 2022; Mareli *et al.*, 2020; Marslen-Wilson *et al.*, 1994; Milin *et al.*, 2018).

Experiment 1

Participants

Participants in Experiment 1 were 48 typical adults, students from several faculties at the Aristotle University of Thessaloniki (24 female, 24 male; 43 right-handed, 5 left-handed; mean age 22.3 (2.46) years; mean year of study 3.9 (1.67) academic years). They were all native speakers of Standard Modern Greek with normal—or corrected to normal—vision and hearing, and no language/reading disorders. Prior to testing, participants gave their informed consent and were offered monetary compensation.

Semantic Transparency Pretests

In order to choose the appropriate items for our priming task, we first conducted three semantic relatedness rating pretests with overall 566 word pairs and 240 independent raters drawn from the same sampling frame as the participants (120 female, 120 male; mean age 20.5 (1.94) years; mean year of study 2.84 (1.53) academic years). Each rater was given a printed list of word pairs and had to rate the semantic relatedness between the members of each word pair using a Likert-type scale, where point 1 stood for minimum and point 7 for maximum semantic

relatedness (Gonnerman *et al.*, 2007; Hill & Gonnerman 2022; Jared *et al.*, 2017; Xu & Taft 2015). Word pairs comprised a suffixed verb or adjective (consisting of a base stem, a derivational suffix, and an inflectional ending; *e.g.*, καρφ-ών-ω /kar'fono/ 'to nail', εθν-ικ-ός /ethni'kos/ 'national') as the first member and its noun base (consisting of the same base stem and an inflectional ending; *e.g.*, καρφ-ί /kar'fi/ 'nail', έθν-ος /ethnos/ 'nation') as the second member. Moreover, some word pairs comprised a pseudo-derived suffixed verb or adjective (with the same surface morphological structure, *i.e.*, base stem + derivational suffix + inflectional ending; *e.g.*, μάλ-ών-ω /bal'ono/ 'to mend', βαλτ-ικ-ός /valti'kos/ 'Baltic') and a putative noun base (consisting of a homonym/homograph base stem and an inflectional ending; *e.g.*, μάλ-α /bala/ 'ball', βάλτ-ος /valtos/ 'swamp'). Participants were asked to judge semantic relatedness regardless of orthographic/phonological similarity, to exploit the whole rating scale, and to circle the "don't know" option (DK) in case they did not know the meaning of a word, instead of answering by chance. Afterwards, for each word pair we calculated the ratings mean (representing semantic transparency), standard deviation (SD), and 95% confidence interval (CI), as well as the percentage of known answers (minus DKs and missing values) representing familiarity.

Materials and Design

From this pool, we selected 90 prime-target pairs, each with a familiarity percentage over 90%, forming three critical semantic transparency conditions (Gonnerman *et al.*, 2007; Jared *et al.*, 2017; Xu & Taft 2015). The transparent condition comprised 30 prime-target pairs, each with a ratings mean and a CI from 5 up to 7. In this condition, primes were constructed complex words sharing with their targets the same base morpheme and a strong semantic relationship, *e.g.*, ζυμώνω – ζύμη /zi'mono/ – /'zimi/ 'to kneed' – 'dough', ratings mean = 6.47 (1.2), CI = [6.12 – 6.82]. The semitransparent condition comprised 30 prime-target pairs, each with a ratings mean and a CI from 3 up to 4.99. In this condition, primes were constructed complex words sharing with their targets the same base morpheme and an intermediate semantic relationship, *e.g.*, ψαρώνω – ψάρι /psa'rono/ – /'psari/ 'to psych out' – 'fish', ratings mean = 3.94 (2), CI = [3.36 – 4.52]. The opaque condition comprised 30 prime-target pairs, each with a ratings mean and a CI from 1 up to 2.99. In this condition, primes were non-constructed complex words sharing with their targets no semantic relationship, but only either an etymological morphological relationship, *e.g.*, σκοτώνω – σκότος /sko'tono/ – /'skotos/ 'to kill' – 'darkness', ratings mean = 1.79 (1.5), CI = [1.36 – 2.23], or an incidental pseudo-morphological relationship, *e.g.*, μαγκώνω – μάγκας /ma'gono/ – /'magas/ 'to grip' – 'buster', ratings mean = 1.48 (1.01), CI = [1.19 – 1.77] (see Longtin *et al.*, 2003; Rastle & Davis 2003; 2008). A one-way analysis of variance (ANOVA) and *post-hoc* Bonferroni comparisons showed significant gradation and no overlap among the three semantic transparency conditions:

transparent, ratings mean = 6.4 (0.299), semitransparent, ratings mean = 4.17 (0.458), opaque, ratings mean = 1.83 (0.423), $F(2, 87) = 984.634$, $p < .001$, $MSE = 0.159$, $\eta^2p = .958$; transparent \neq semitransparent, $p < .001$, semitransparent \neq opaque, $p < .001$, transparent \neq opaque, $p < .001$.¹

In order to minimize confounding effects, primes and targets were groupwise matched across the three semantic transparency conditions for several extraneous variables known to influence lexical and morphological processing, such as word length (number of letters, graphemes, phones, phonemes, and syllables), word and lemma frequency, syllable and bigram frequency, morphological family size, prime-target formal (orthographic, graphemic, phonetic, and phonological) similarity, grapheme-phoneme transparency, orthographic and phonological neighborhood size (Levenshtein Distance 20), orthographic and phonological uniqueness point. Primes were also matched for part of speech, derivational suffix, and inflectional ending, and noun targets were further matched for the semantic feature [concrete].

Control primes were morphologically, formally (orthographically/phonologically), and semantically unrelated to the targets, though pairwise matched to the corresponding critical primes for the above-mentioned extraneous variables—with the exception of prime-target formal similarity (*e.g.*, target ζύμη / 'zimi/ 'dough', critical prime ζυμώνω /zi'mono/ 'to kneed', control prime θολώνω /θo'lono/ 'to blur'). Critical and control primes were divided into two counterbalancing lists, each containing 45 critical and 45 control prime-target pairs (15 per semantic transparency condition). Half participants were assigned to List 1 and the other half to List 2, so that every participant was exposed to all priming conditions, never seeing the same target twice. Finally, in each list we added 90 word-word fillers and 180 word-nonword fillers. By doing so, on the one hand we balanced the targets for lexical status (180 words, 180 nonwords) and the primes for morphological complexity (180 suffixed primes, 180 unsuffixed primes), on the other hand we lowered the prime-target relatedness proportion to 25 % (90 related prime-target pairs, 270 unrelated prime-target pairs), so as to diminish potential strategic effects due to the conscious perception of the prime in overt priming (Milin *et al.*, 2018, 244-245; Napps & Fowler 1987).

Apparatus and Procedure

The experiment was implemented in E-Prime 2.0 (Schneider *et al.*, 2012) using an Acer laptop (LCD screen, 15.6 inches, 1,366 x 768 pixels, 60 Hz). Stimuli were presented in the center of the screen, in a black font against a white background, according to an unmasked visual priming protocol as follows: fixation cross for 500 ms; prime in lowercase letters, Times New Roman, 30 pt, bold, for 250 ms;

1. All statistical tests reported are two-tailed with the significance level (α) set at .05.

target in lowercase letters, Times New Roman, 40 pt, bold, until response;¹ blank screen (inter-trial interval) for 1,000 ms. Participants were individually tested in a quiet, dimly-lit room, seated at a distance of about 60 cm from the laptop screen. In the instructions, we asked them to make lexical decisions to the targets by pressing the ‘/?’ key or the ‘Z’ key on the keyboard with their right and left index fingers, the dominant hand always controlling the ‘word’ response. Participants were, also, prior informed that a prime would appear in the center of the screen before the target and were requested to respond as quickly and as accurately as possible. An individual session, lasting approximately 20 minutes, comprised the instructions, 16 practice trials, and 360 experimental trials presented in a different randomized order, with three self-administered breaks (one break every 90 experimental trials).

Results and Discussion

We analyzed a total of 4,026 (95.31%) RT points, after removing: two targets from the semitransparent condition because of their low accuracy mean (< 70 %); 127 (3.01%) RT points for incorrect responses; and 71 (1.68 %) outliers greater than three SDs above the mean. ANOVAs by participants (*F*₁) and by items (*F*₂) were conducted on RT (measured in ms), entering prime type (critical, control) as a factor within participants and within items, and semantic transparency (transparent, semitransparent, opaque) as a factor within participants and between items. Table 1 summarizes the RT mean (SD) and priming effect per condition, by participants and items.

Table 1—Descriptive statistics for RTs (ms) by participants/items, in Experiment 1

	Transparent	Semitransparent	Opaque
Control	610 (89.8) / 612 (47.8)	634 (85) / 639 (71.4)	613 (83.8) / 613 (55.2)
Critical	581 (103) / 579 (31.2)	598 (99) / 599 (47.7)	616 (88.2) / 617 (45.6)
Priming Effect	29 / 33	36 / 40	-3 / -4

ANOVAs showed a statistically significant main effect of prime type: *F*₁(1, 47) = 30.275, *p* < .001, MSE = 1,034.310, η^2p = .392; *F*₂(1, 85) = 20.930, *p* < .001, MSE = 1,109.576, η^2p = .198; and a significant interaction between prime type and semantic transparency: *F*₁ (2, 94) = 11.208, *p* < .001, MSE = 930.627, η^2p = .193; *F*₂(2, 85) = 7.041, *p* = .001, MSE = 1,109.576, η^2p = .142. Follow-up paired *t*-tests revealed significant priming in the transparent condition: *t*₁(47) = 4.760, *p* < .001; *t*₂(29) = 4.064, *p* < .001; and the semitransparent condition: *t*₁(47) = 5.129, *p* <

1. The target is presented in lowercase letters, so as to bear the word stress diacritic, and in a larger font size than the prime, so as to be perceptually distinct from the prime (see ORFANIDOU *et al.*, 2011; TSAPKINI *et al.*, 2002; VOGA & GRAINGER 2004).

.001; $t_2(27) = 3.588$, $p = .001$; but insignificant negative priming in the opaque condition: $t_1(47) = -0.494$, $p = .624$; $t_2(29) = -0.490$, $p = .628$. Moreover, *post-hoc* Bonferroni comparisons revealed no significant difference of the priming effect between the transparent and semitransparent condition ($p_1 = 1$; $p_2 = 1$), but a significant difference between the transparent and opaque condition ($p_1 = .002$; $p_2 = .011$), as well as between the semitransparent and opaque condition ($p_1 < .001$; $p_2 = .003$). That is, the priming effect was modulated by semantic transparency in a dual manner: statistically significant priming of the same magnitude in the transparent and semitransparent condition, yet no priming at all in the opaque condition.

To further test for a potential graded effect, we conducted simple linear regression analyses by items, entering semantic transparency as a numeric predictor (*i.e.*, the semantic transparency ratings mean of each item taken from the pretests) and priming as the dependent variable. We found a statistically significant effect of semantic transparency in total: $F(1, 86) = 11.072$, $p = .001$, $R^2 = .114$; but neither for each category of items analyzed apart: transparent, $F(1, 28) = 2.157$, $p = .153$, $R^2 = .072$; semitransparent, $F(1, 26) = 0.603$, $p = .445$, $R^2 = .023$; opaque, $F(1, 28) = 1.951$, $p = .173$, $R^2 = .065$; nor for the transparent and semitransparent items analyzed together: $F(1, 56) < 0.001$, $p = .989$, $R^2 < .001$. So, it seems that semantic transparency exerts a significant influence on priming, but this influence is not graded, as it was found in the ANOVAs.

In addition, to corroborate the findings from the ANOVAs and the simple linear regression analyses, we carried out corresponding analyses of covariance (ANCOVAs) and multiple linear regression analyses, entering the controlled extraneous variables (*e.g.*, word length, word/lemma frequency, morphological family size, prime-target formal similarity, etc.) as numeric covariates and predictors, respectively. ANCOVAs and multiple linear regression analyses verified the initial findings, showing no significant effects of the extraneous variables.

Finally, Table 2 summarizes the error rates mean (SD) and priming effect per condition, by participants and items. Error rates were very low. ANOVAs by participants/items and simple linear regression analyses by items were also conducted on error rates, yielding no significant effects. Moreover, we found no speed-accuracy trade-off, since RTs and error rates were not significantly correlated by participants ($r_1 = -.125$, $p = .397$) and were positively correlated by items ($r_2 = .606$, $p < .001$).

Table 2—Descriptive statistics for error rates (%) by participants/items, in Experiment 1

	Transparent	Semitransparent	Opaque
Control	3.27 (4.37) / 3.1 (6.17)	3.65 (5.2) / 3.68 (7.54)	3.52 (5.84) / 3.37 (7.29)
Critical	0.87 (2.34) / 0.8 (1.94)	3.5 (4.57) / 3.5 (6.92)	3.35 (4.91) / 3.27 (5.18)
Priming Effect	2.4 / 2.3	0.15 / 0.18	0.17 / 0.1

The results of Experiment 1 revealed that semantic transparency significantly affected morphological priming in a dual manner. Unmasked visual priming was statistically significant and equivalent between the transparent and semitransparent condition, though absent in the opaque condition (transparent = semitransparent > opaque). This dual priming pattern can be interpreted assuming constructed representations for transparent and semitransparent derived words, and non-constructed representations for opaque derived words, in line with the associative construction perspective. Thus, it appears that the observed priming effect stems from the shared prime-target base morpheme that is recognized as a morphosemantic unit during the processing of constructed words, unlike the processing of non-constructed words. On the other hand, a dissociative account of morphology would predict either a flat priming pattern (transparent = semitransparent = opaque) originating solely from surface morphological complexity or a graded priming pattern (transparent > semitransparent > opaque) originating solely from semantics, neither of which is attested in our experimental findings.

Experiment 2

Participants

Forty-eight students from several faculties at the Aristotle University of Thessaloniki (24 female, 24 male; 44 right-handed, 4 left-handed; mean age 21.8 (3.12) years; mean year of study 3.38 (1.77) academic years) participated in Experiment 2. They were all typical adults, native speakers of Standard Modern Greek with normal—or corrected to normal—vision and hearing, and they reported no language or reading disorders. All participants gave their informed consent prior to testing and were offered monetary compensation. None of them took part in Experiment 1 or the in the semantic transparency pretests.

Materials and Design

The design and materials were the same as in Experiment 1.

Apparatus and Procedure

Experiment 2 was implemented by means of E-Prime 2.0 and an Acer laptop, exactly as Experiment 1. Auditory files (WAV, 44.1 KHz, 16 bit, mono) were recorded and edited on the same laptop using Praat (version 6.1.09) and a Blue Yeti USB microphone, and they were displayed through a pair of Microsoft USB headphones. Visual stimuli appeared in the center of the screen, in a black font against a white background. The sequence of the events in a trial was as follows: fixation cross for 500 ms; auditory prime with a blank screen; visual target immediately after the prime offset (SOA = auditory prime duration) in lowercase letters, Times New Roman, 40 pt, bold, until response; blank screen (inter-trial interval) for 1,000 ms. Procedure, instructions, and session details were the same as in Experiment 1, except that participants had to make lexical decisions to the targets after listening to the primes.

Results and Discussion

We analyzed a total of 4,021 (95.2 %) RT points, after dismissing: two targets from the semitransparent condition due to their low accuracy mean ($< 70\%$); 134 (3.17 %) RT points for incorrect responses; and 69 (1.63 %) outliers greater than three SDs above the mean. ANOVAs by participants (F_1) and by items (F_2) were conducted on RT, entering prime type (critical, control) as a factor within participants and within items, and semantic transparency (transparent, semitransparent, opaque) as a factor within participants and between items. The RT mean (SD) and priming effect per condition are summarized in Table 3.

Table 3—Descriptive Statistics for RTs (ms) by Participants/Items, in Experiment 2

	Transparent	Semitransparent	Opaque
Control	615 (84.1) / 617 (58.8)	634 (82.4) / 637 (67.4)	614 (75.7) / 617 (56.6)
Critical	578 (81.4) / 577 (27)	605 (84.7) / 607 (54)	618 (90.7) / 619 (49.3)
Priming Effect	37 / 40	29 / 30	-4 / -2

We found a statistically significant main effect of prime type: $F_1(1, 47) = 22.804$, $p < .001$, $MSE = 1,324.762$, $\eta^2p = .327$; $F_2(1, 85) = 20.178$, $p < .001$, $MSE = 1,090.761$, $\eta^2p = .192$; and a significant interaction between prime type and semantic transparency: $F_1(2, 94) = 10.017$, $p < .001$, $MSE = 1,102.482$, $\eta^2p = .176$; $F_2(2, 85) = 6.635$, $p = .002$, $MSE = 1,090.761$, $\eta^2p = .135$. Follow-up paired t -tests showed significant priming in the transparent condition: $t_1(47) = 6.810$, $p < .001$; $t_2(29) = 5.322$, $p < .001$; and the semitransparent condition: $t_1(47) = 3.775$, $p < .001$; $t_2(27) = 3.082$, $p = .005$; yet insignificant negative priming in the opaque condition: $t_1(47) = -0.487$, $p = .628$; $t_2(29) = -0.262$, $p = .795$. *Post-hoc* Bonferroni comparisons revealed no significant difference of the priming effect between the transparent and semitransparent condition ($p_1 = .969$; $p_2 = 1$), though a significant difference between the transparent and opaque condition ($p_1 < .001$; $p_2 = .002$), as well as between the semitransparent and opaque condition ($p_1 = .010$; $p_2 = .034$). Again, as in Experiment 1, we observe a dichotomous semantic transparency effect; the magnitude of priming is statistically significant and equivalent in the transparent and semitransparent condition, but negligible in the opaque condition.

Furthermore, simple linear regression analyses by items, with semantic transparency as a numeric predictor and priming as the dependent variable, revealed a statistically significant effect of semantic transparency in total: $F(1, 86) = 13.222$, $p < .001$, $R^2 = .133$; but neither for each category of items analyzed apart: transparent, $F(1, 28) = 1.355$, $p = .254$, $R^2 = .046$; semitransparent, $F(1, 26) = 0.543$, $p = .468$, $R^2 = .020$; opaque, $F(1, 28) = 0.045$, $p = .834$, $R^2 = .002$; nor for the transparent and semitransparent items analyzed together: $F(1, 56) = 1.506$, $p = .225$, $R^2 = .026$. Namely, the effect of semantic transparency as a numeric predictor was significant but not graded, in accordance with the ANOVAs.

These findings were, additionally, confirmed by corresponding ANCOVAs and multiple linear regression analyses yielding no significant effects of the extraneous variables entered as numeric covariates/predictors. Finally, the error rates mean (SD) and priming effect per condition are summarized in Table 4. Error rates were very low. ANOVAs and simple linear regression analyses on error rates yielded similar results with the corresponding RT analyses. Moreover, the correlation between RTs and error rates was not significant by participants ($r_1 = -.246$, $p = .092$) and was significantly positive by items ($r_2 = .669$, $p < .001$) revealing no speed-accuracy trade-off.

Table 4—Descriptive Statistics for Error Rates (%) by Participants/Items, in Experiment 2

	Transparent	Semitransparent	Opaque
Control	4.13 (5.31) / 3.93 (7.3)	3.65 (5.4) / 3.64 (7.73)	3.52 (5.5) / 3.33 (4.59)
Critical	1.15 (2.92) / 1.07 (2.08)	2.63 (4.48) / 2.64 (5.57)	4.08 (4.88) / 4 (6.38)
Priming Effect	2.98 / 2.86	1.02 / 1	-0.56 / -0.67

The results of Experiment 2 indicate a dichotomous semantic transparency effect on morphological priming, consistent with the results of Experiment 1. Immediate cross-modal priming was statistically significant and equivalent between the transparent and semitransparent prime-target pairs, but almost null for the opaque prime-target pairs (transparent = semitransparent > opaque). These findings fit well within the associative construction approach, supposing that transparent and semitransparent derived words are represented as both morphologically complex and constructed, while opaque derived words are represented as morphologically complex but non-constructed. Therefore, the base morpheme functions as a morphosemantic representational unit in both the transparent and the semitransparent derived words (*i.e.*, constructed words), inducing priming of the same magnitude. On the other hand, the base morpheme induces no significant opaque priming, since it does not function as a morphosemantic representational unit in the opaque derived words (*i.e.*, non-constructed words). In contrast, the dissociative approach cannot easily account for this dichotomous priming pattern, expecting either a graded semantic effect (transparent > semitransparent > opaque) or a comparable formal effect (transparent = semitransparent = opaque) among the critical conditions.

General Discussion and Conclusions

In this study, we attempted to experimentally test the associative construction approach to morphological theory, according to which constructed words share not only form but also meaning with their constituent morphemes, in contradis-

tion to the dissociative approach to morphological theory, according to which the structure of a morphologically complex word is independent of semantics. For this purpose, we investigated whether meaning affects the representation and processing of derived words in Modern Greek by means of two lexical decision experiments with overt morphological priming, conducted with typical adults. The experimental design comprised three critical semantic transparency conditions (transparent, semitransparent, opaque) and was applied with an unmasked visual priming protocol (Experiment 1) and an immediate cross-modal priming protocol (Experiment 2), considered to tap into the later morphological processing and the central morphological representation in the core of the mental lexicon. In both experiments, transparent and semitransparent priming was statistically significant, though opaque priming was statistically insignificant and numerically negligible. Moreover, the magnitude of priming was statistically equivalent between the transparent and semitransparent condition. This dual priming pattern (transparent = semitransparent > opaque) indicates that the observed effect is neither purely formal nor purely semantic, as expected by the dissociative account, but rather morphosemantic, as predicted by the associative account. Therefore, we can give a 'yes' answer to our research question, suggesting that the mental representation and processing of morphological structure is critically affected by meaning.

Our findings underpin an associative account of derivational morphology in Modern Greek, treating the semantically transparent and semitransparent complex words as constructed words, and the semantically opaque complex (or pseudo-complex) words as non-constructed words. From this point of view, the locus of priming appears to be the shared prime-target base morpheme that is represented and processed as a genuine morpheme (*i.e.*, a unit combining both form and meaning) in the case of constructed words (*i.e.*, transparent and semitransparent complex words), but not in the case of non-constructed words (*i.e.*, opaque complex and pseudo-complex words). Thus, it can be reasonably explained why transparent and semitransparent derived words facilitate to the same extent the recognition of their noun bases, unlike the opaque derived and pseudo-derived words that do not facilitate at all the recognition of their etymological or putative noun bases. Instead, a dissociative account of derivational morphology could not readily accommodate either the equal magnitude of transparent and semitransparent priming or the total absence of opaque priming in our experimental data.

It is, also, worth mentioning that these two opposite theoretical approaches to morphology align with different psycholinguistic accounts of morphological processing. Namely, the associative construction approach is consistent with a supralexical account of morphological processing considering morphemes as morphosemantic representations mediating between lexical representations and semantic representations (Giraudo & Grainger 2001; 2003; Voga & Giraudo 2017). On the other hand, the dissociative approach ties in with a full-decomposition account of morphological processing postulating the compositional repre-

sensation of all morphologically complex words irrespective of meaning (Fruchter & Marantz 2015; Marantz 2013; Smolka *et al.*, 2014), as well as with a strong version of sublexical decomposition treating morphemes as morpho-orthographic (or morpho-phonological) representations confined at a level under/before the level of lexical representations (Crepaldi *et al.*, 2010; Longtin & Meunier 2005; Rastle & Davis 2008). Consequently, our findings supporting the associative construction approach feature morphemes as morphosemantic representations in the mental lexicon.

To conclude, this study provides experimental evidence arguing for an associative construction approach to derivational morphology in Modern Greek. Our results imply that semantically transparent and semitransparent complex words are constructed words with internal morphological structure, while semantically opaque complex words are non-constructed words without internal morphological structure. Considering that the representation and processing of complex words is semantically constrained, we suggest that morphological analysis must take both form and meaning into account, rather than being solely form-based and semantically blind.

Acknowledgments

I wish to thank Professor Emerita Anna Anastassiadis-Symeonidis and Professor Despina Papadopoulou for their helpful comments on an earlier draft of this paper.

References

- ACQUAVIVA, Paolo, “Roots and lexicality in distributed morphology”, *York Papers in Linguistics Series 2*, 10 (2009): 1–21.
- AMENTA, Simona, & CREPALDI, Davide, “Morphological processing as we know it: An analytical review of morphological effects in visual word identification”, *Frontiers in Psychology* 3 (2012): Article 232: <https://doi.org/10.3389/fpsyg.2012.00232>.
- ANASTASSIADIS-SYMEONIDIS Anna ΑΝΑΣΤΑΣΙΑΔΗ-ΣΥΜΕΩΝΙΔΗ Άννα, “I neoelliniki paragogi kata to montelo tis D. Corbin Η νεοελληνική παραγωγή κατά το μοντέλο της D. Corbin” [Modern Greek derivation within D. Corbin’s model], *Meletes gia tin Elliniki Glossa Μελέτες για την Ελληνική Γλώσσα* [Studies in Greek Linguistics], 13 (1992): 505–526.
- ANASTASSIADIS-SYMEONIDIS, Anna, “On Modern Greek denominal adjectives”, in G. E. BOOIJ, A. RALLI & S. SCALISE (ed.), *Proceedings of the 1st Mediterranean Morphology Meeting: Allomorphy, Compounding, Inflection*, Patras: University of Patras, 1998, 29–40.

- ANASTASSIADIS-SYMEONIDIS, Anna, "Suffixation and the expression of time and space in Modern Greek", *Lingue e Linguaggio* 15.1 (2016): 59–84.
- ANASTASSIADIS-SYMEONIDIS, Anna, "Arches morfologikis analysis me efarmogi stin epithimato poiisi Αρχές μορφολογικής ανάλυσης με εφαρμογή στην επιθηματοποίηση" [Principles of morphological analysis applying to suffixation], *Meletes gia tin Elliniki Glossa Μελέτες για την Ελληνική Γλώσσα* [Studies in Greek Linguistics], 42 (2023): 13–19.
- ANASTASSIADIS-SYMEONIDIS, Anna, *I epithimato poiisi sti nea elliniki H επιθηματοποίηση στη νέα ελληνική* [Suffixation in Modern Greek], forthcoming.
- ARONOFF, Mark, *Word Formation in Generative Grammar*, Cambridge Massachusetts: The MIT Press, 1976.
- BAUER, Laurie, Livia KÖRTVÉLYESSY & Pavol ŠTEKAUER (ed.), *Semantics of Complex Words*, Cham: Springer, 2015.
- CHOMSKY, Noam, "Remarks on nominalization", in R. A. JACOBS & P. S. ROSENBAUM (ed.), *Readings in English Transformational Grammar*, Waltham Massachusetts: Ginn and Company, 1970, 184–221.
- CHOMSKY, Noam & HALLE, Morris, *The Sound Pattern of English*, New York: Harper and Row, 1968.
- CORBIN, Danielle, *Morphologie dérivationnelle et structuration du lexique*, Tübingen / Villeneuve d'Ascq: Max Niemeyer Verlag / Presses universitaires de Lille, 1987/1991.
- CORBIN, Danielle, "Form, structure and meaning of constructed words in an associative and stratified lexical component", in G. E. BOOIJ & J. VAN MARLE (ed.), *Yearbook of Morphology Volume 2*, Dordrecht: Foris, 1989, 31–54.
- CORBIN, Danielle, "Introduction. La formation des mots: Structures et interprétations", *Lexique* 10 (1991): 7–30.
- CORBIN, Danielle, *Le lexique construit. Méthodologie d'analyse*, forthcoming.
- CREPALDI, Davide, Kathleen RASTLE, Max COLTHEART & Lyndsey NICKELS, "'Fell' primes 'fall', but does 'bell' prime 'ball'? Masked priming with irregularly-inflected primes", *Journal of Memory and Language* 63.1 (2010): 83–99.
- Dictionary of Standard Modern Greek*, Thessaloniki: Institute of Modern Greek Studies, 1998.
- DIEPENDAELE, Kevin, Jonathan GRAINGER & Dominiek SANDRA, "Derivational morphology and skilled reading: An empirical overview", in M. J. SPIVEY, K. McRAE & M. F. JOANISSE (ed.), *The Cambridge Handbook of Psycholinguistics*, Cambridge: Cambridge University Press, 2012, 311–332.

- DI SCIULLO, Anna Maria & WILLIAMS, Edwin, *On the Definition of Word*, Cambridge Massachusetts: The MIT Press, 1987.
- EFTHYMIΟΥ, Angeliki, *Le suffixe -ιά en grec moderne: La manifestation d'un degré maximal d'anthropocentricité*, Lille / Saarbrücken, Presses universitaires du Septentrion / Éditions universitaires européennes, 1999/2012.
- EFTHYMIΟΥ Angeliki, *O schimatismos ton rimaton sti nea elliniki glossa O σχηματισμός των ρημάτων στη νέα ελληνική γλώσσα* [Verb formation in Modern Greek], Thessaloniki: Epikentro, 2018.
- EFTHYMIΟΥ Angeliki, "Paragogi Παράγωγή" [Derivation], in D. PAPADOPOULOU & A. REVITHIADOU (ed.), *Eisagogi sti morfologia: Theoria kai peiramatikes efarmoges* Εισαγωγή στη μορφολογία: Θεωρία και πειραματικές εφαρμογές [Introduction to morphology: Theory and experimental applications], Thessaloniki: Institute of Modern Greek Studies, 2023, 199–234.
- EFTHYMIΟΥ, Angeliki, Georgia FRAGAKI & Angelos MARKOS, "Productivity of verb-forming suffixes in Modern Greek: A corpus-based study", *Morphology* 22.4 (2012): 515–543.
- FELDMAN, Laurie B. & Petar MILIN, "Psycholinguistic studies of word morphology and their implications for models of the mental lexicon and lexical processing", in R. BERTHIAUME, D. DAIGLE & A. DESROCHERS (ed.), *Morphological Processing and Literacy Development: Current Issues and Research*, New York: Routledge Taylor and Francis Group, 2018, 16–47.
- FELDMAN, Laurie B. & Fermín MOSCOSO DEL PRADO MARTÍN, "Tuning language processing mechanisms to a language's morphology without decomposition: The case of semantic transparency", in A. D. SIMS, A. USSISHKIN, J. PARKER & S. WRAY (ed.), *Morphological Diversity and Linguistic Cognition*, Cambridge: Cambridge University Press, 2022, 31–55.
- FRUCHTER, Joseph & Alec MARANTZ, "Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition", *Brain and Language* 143 (2015): 81–96.
- GIRAUDO, Hélène & Jonathan GRAINGER, "Priming complex words: Evidence for supralelexical representation of morphology", *Psychonomic Bulletin and Review* 8.1 (2001): 127–131.
- GIRAUDO, Hélène & Jonathan GRAINGER, "A supralelexical model for French derivational morphology", in E. M. H. ASSINK & D. SANDRA (ed.), *Reading Complex Words: Cross-Language Studies*, New York: Kluwer Academic / Plenum Publishers, 2003, 139–157.
- GONNERMAN, Laura M., Mark S. SEIDENBERG & Elaine S. ANDERSEN, "Graded semantic and phonological similarity effects in priming: Evidence for a

- distributed connectionist approach to morphology”, *Journal of Experimental Psychology: General* 136.2 (2007): 323–345.
- HALLE, Morris & Alec MARANTZ, “Distributed morphology and the pieces of inflection”, in K. HALLE & S. J. KEYSER (ed.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge Massachusetts: The MIT Press, 1993, 111–176.
- HILL, Katherine J. & Laura M. GONNERMAN, “Graded morphological processing in French”, *Lexique* 30 (2022): 57–79.
- HOCKETT, Charles F., *A Course in Modern Linguistics*, New York: Macmillan, 1958.
- JACKENDOFF, Ray, “Morphological and semantic regularities in the lexicon”, *Language* 51.3 (1975): 639–671.
- JARED, Debra, Olessia JOURAVLEV & Marc F. JOANISSE, “The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials”, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43.3 (2017): 422–450.
- KIELAR, Aneta & Marc F. JOANISSE, “The role of semantic and phonological factors in word recognition: An ERP cross-modal priming study of derivational morphology”, *Neuropsychologia* 49.2 (2011): 161–177.
- KOTOWSKI, Sven & Ingo PLAG, (ed.), *The Semantics of Derivational Morphology: Theory, Methods, Evidence*, Berlin / Boston: De Gruyter, 2023.
- LIEBER, Rochelle, *On the Organization of the Lexicon*, Bloomington: Indiana University Linguistics Club, 1981.
- LIEBER, Rochelle, *Deconstructing Morphology: Word Formation in Syntactic Theory*, Chicago: University of Chicago Press, 1992.
- LIEBER, Rochelle, *Morphology and Lexical Semantics*, Cambridge: Cambridge University Press, 2004.
- LONGTIN, Catherine-Marie & Fanny MEUNIER, “Morphological decomposition in early visual word processing”, *Journal of Memory and Language* 53.1 (2005): 26–41.
- LONGTIN, Catherine-Marie, Juan SEGUI & Pierre A. HALLÉ, “Morphological priming without morphological relationship”, *Language and Cognitive Processes* 18.3 (2003): 313–334.
- MARANTZ, Alec, “No escape from morphemes in morphological processing”, *Language and Cognitive Processes* 28.7 (2013): 905–916.
- MARELLI, Marco, Daniela TRAFICANTE & Cristina BURANI, “Reading morphologically complex words: Experimental evidence and learning models”, in V.

- PIRRELLI, I. PLAG & W. U. DRESSLER (ed.), *Word Knowledge and Word Usage: A Cross-Disciplinary Guide to the Mental Lexicon*, Berlin / Boston: De Gruyter Mouton, 2020, 553–592.
- MARSLÉN-WILSON, William D., Mirjana BOZIC & Billi RANDALL, “Early decomposition in visual word recognition: Dissociating morphology, form, and meaning”, *Language and Cognitive Processes* 23.3 (2008): 394–421.
- MARSLÉN-WILSON, William D., Lorraine K. TYLER, Rachelle WAKSLER & Lianne OLDER, “Morphology and meaning in the English mental lexicon”, *Psychological Review* 101.1 (1994): 3–33.
- MILIN, Petar, Eva SMOLKA & Laurie B. FELDMAN, “Models of lexical access and morphological processing”, in E. M. FERNÁNDEZ & H. SMITH CAIRNS (ed.), *The Handbook of Psycholinguistics*, Hoboken New Jersey: Wiley Blackwell, 2018, 240–268.
- NAPPS, Shirely E. & Carol A. FOWLER, “Formal relationships among words and the organization of the mental lexicon”, *Journal of Psycholinguistic Research* 16.3 (1987): 257–272.
- ORFANIDOU, Eleni, Matthew H. DAVIS & William D. MARSLÉN-WILSON, “Orthographic and semantic opacity in masked and delayed priming: Evidence from Greek”, *Language and Cognitive Processes* 26.4–6 (2011): 530–557.
- PANAGIOTIDIS, Phoebos, “On the nature of roots: Content, form, identification”, *Evolutionary Linguistic Theory* 2.1 (2020), 56–83.
- PANAGIOTIDIS Phoebos, *Mesa apo tis lexeis Μέσα από τις λέξεις* [Through words], Athens: Ekdoseis tou eikostou protou, 2021.
- PANAGIOTIDIS Phoebos, *Metaxy noisis kai fonis Μεταξύ νόησης και φωνής* [Between cognition and voice], Athens: Nisos, 2023.
- PANAGIOTIDIS, Phoebos, Vassilios SPYROPOULOS & Anthi REVITHIADOU, “Little v as a categorizing verbal head: Evidence from Greek”, in R. D’ALESSANDRO, I. FRANCO & Á. J. GALLEGÓ (ed.), *The Verbal Domain*, Oxford: Oxford University Press, 2017, 29–48.
- RAINER, Franz, Francesco GARDANI, Hans Christian LUSCHÜTZKY & Wolfgang U. DRESSLER (ed.), *Morphology and Meaning: Selected Papers from the 15th International Morphology Meeting*, Amsterdam / Philadelphia: John Benjamins, 2014.
- RALLI Angeliki, *Morfologia: Nea, anatheorimeni ekdosi Μορφολογία: Νέα, αναθεωρημένη έκδοση* [Morphology: New, revised edition], Athens: Patakis, 2022.

- RASTLE, Kathleen & Matthew H. DAVIS, "Reading morphologically complex words: Some thoughts from masked priming", in S. KINOSHITA & S. J. LUPKER (ed.), *Masked Priming: The State of the Art*, New York / Hove: Psychology Press, 2003, 279–305.
- RASTLE, Kathleen & Matthew H. DAVIS, "Morphological decomposition based on the analysis of orthography", *Language and Cognitive Processes* 23.7-8 (2008): 942–971.
- SCHNEIDER, Walter, Amy ESCHMAN & Anthony ZUCCOLOTTO, *E-Prime User's Guide*, Pittsburgh: Psychology Software Tools, 2012.
- SELKIRK, Elisabeth O., *The Syntax of Words*, Cambridge Massachusetts: The MIT Press, 1982.
- SMOLKA, Eva, Katrin H. PRELLER & Carsten EULITZ, "'Verstehen' ('understand') primes 'stehen' ('stand'): Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs", *Journal of Memory and Language* 72 (2014): 16–36.
- TOMAN, Jindrich, *Wortsyntax: Eine Diskussion ausgewählter Probleme deutscher Wortbildung*, Tübingen: Max Niemeyer Verlag, 1987.
- TOMAN, Jindrich, "Word syntax", in A. SPENCER & A. M. ZWICKY (ed.), *The Handbook of Morphology*, Oxford / Malden: Blackwell, 1998, 306–321.
- TSAPKINI, Kyrana, Gonia JAREMA & Eva KEHAYIA, "Regularity revisited: Evidence from lexical access of verbs and nouns in Greek", *Brain and Language* 81.1–3 (2002): 103–119.
- VOGA, Madeleine & Hélène GIRAUDO, "Word and beyond-word issues in morphological processing", *Word Structure* 10.2 (2017): 235–255.
- VOGA, Madeleine & Jonathan GRAINGER, "Masked morphological priming with varying levels of form overlap: Evidence from Greek verbs", *Current Psychology Letters: Behaviour, Brain, and Cognition* 13.2 (2004): Article 160504, <https://doi.org/10.4000/cpl.422>.
- XU, Joe & Marcus TAFT, "The effects of semantic transparency and base frequency on the recognition of English complex words", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41.3 (2015): 904–910.

L'étymologie populaire comme mécanisme de changement linguistique : apport des données du grec et implications pour le lexique mental

Assimakis FLIATOURAS, Madeleine VOGA
& Anna ANASTASSIADIS-SYMÉONIDIS

*Université Démocrite de Thrace, Grèce
ReSO, Université de Montpellier Paul-Valéry
Université Aristote de Thessalonique, Grèce*

Introduction¹

Ce travail se donne plusieurs objectifs : tout d'abord, définir et classer les différents types d'étymologie populaire sur la base de l'acception qu'il s'agit d'un processus psycholinguistique. Nous fournirons des preuves en faveur de l'idée que les processus liés à l'étymologie populaire sont régis par des mécanismes systématiques que l'on pourrait rapprocher à des règles, observables au niveau de l'usage. Ces processus sont non pas uniquement de nature phonologique, mais aussi morphologique, comme nous le montrerons à travers de nombreux exemples. *In fine*, notre contribution vise à spécifier le mécanisme lié à l'étymologie populaire ainsi qu'à déceler son fonctionnement d'un point de vue phonologique, morphologique et sémantique. Les données présentées ici sont majoritairement tirées du *Dictionnaire du Grec Moderne* de la Fondation Manolis Triantafyllidis (Petrounias 1998).

1. Abréviations : ALL : allemand, angl. : anglais, FR : français, GA : grec ancien, GH : grec hellénistique/koiné, GMéd. : grec médiéval, GM : grec moderne, IT : italien, LT : latin, NL : néo-latin, PS : persan, SL : slave, TC : turc, VN : vénitien, n. : nom, adj. : adjectif, UL : unité lexicale. La translittération des exemples grecs suit la périodisation de la langue (pour GA ω = o; ; η = e; ; χ = kh, θ = th).

Définition de l'étymologie populaire

L'étymologie est la branche de linguistique historique qui a comme objet d'étude l'origine, l'histoire et l'évolution d'une unité lexicale. L'étymologie populaire¹ correspond à un rapprochement étymologique erronée d'une unité lexicale à une autre, ayant comme effet un changement de forme et/ou de sens² dans le but d'accroître la transparence morphosémantique pour le locuteur (cf. Förstermann 1852, 2 ; Palmer 1882, 250 ; Antilla 1977, 92 ; Baldinger 1973, 241-7 ; Mc-Mahon [1994] 2005, 118 ; Bauer 2006 ; Liberman 2009 ; Müller *et al.* 2015, et pour le grec, Hatzidakis 1915-6, 172-3 ; Makri-Tsilipakou 1986 ; Syméonidis 1992 ; Setatos 1994, 94 ; Anastassiadis-Syméonidis 2003a ; Karantzola & Fliatouras 2004, 94 ; Moïsiadis 2005, 250-4 ; Papanastassiou 2008, 205-212 ; Fliatouras 2018, 136-160). Par ex. (a) *yiuvá'rlaki* < turc *yuvarlak* « boulette de viande et riz » est ainsi rapproché à *yiuvare'laki*, à partir du mot *va'reli* « tonneau » et (b) *e'yoklima* « chèvre-feuille » qui a comme premier constituant le thème du GA *aiks*, *ai'gos*³ « chèvre » est rapproché au mot *'ayios* « saint » :

(1)

(a)

TC	yu	varl	ak	→ GM <i>yiuvá'rlaki</i> 'boulette de viande'
		↓		
GM		va'reli 'tonneau'		

(b)

GMéd.	eyo 'chèvr-'	'klíma 'vigne'	→ GMéd. a'yíoklíma 'chèvrefeuille'
	↓		
GMéd.	'ayios 'saint'		

Enfin, le terme « étymologie populaire », largement répandu, est plutôt réducteur, puisque la participation d'un niveau de langue populaire dans cette procédure n'est pas une règle mais une tendance forte : comme nous allons le voir plus loin, il y a assez d'exemples en GM pour montrer que l'étymologie populaire conduit parfois à des formes davantage savantes, résultant d'une sur-correction ou d'une intervention savante sur une forme plus familière (cf. aussi Béguelin 2002⁴) :

1. Comme nous le verrons plus loin, la « fausse étymologie », due à une erreur étymologique, se différencie de l'étymologie populaire dans la mesure où elle n'affecte pas l'histoire et le changement de l'unité lexicale.

2. Comme nous le verrons plus loin, l'étymologie populaire affecte le plus souvent la forme et le sens et plus rarement, uniquement la forme ou le sens (MOÏSIADIS 2005 : 253-4).

3. Prononciation en GM [eks, e'γos].

4. Pour les fausses étymologies volontaires des savants de l'antiquité et du Moyen Âge, cf. BRUCKER (1988 : 62-6).

- | | | |
|-----|-----|--|
| (2) | (a) | IT <i>elevatore</i> > GM <i>eleva'tori</i> > <i>anava'tori(o)</i> 'monte-charges, ascenseur' selon GA <i>ana'baino</i> ¹ : 'monter' |
| | (b) | FR <i>Gavroche</i> 'héros de V. Hugo' > <i>gavroche</i> > GM <i>γavri'as</i> selon GA <i>gauri'o</i> ² : 'avoir l'air fier' |
| | (c) | FR <i>théodolite</i> > GM <i>θεο'δολιχος</i> selon GA <i>doli'khos</i> ³ 'long' |
| | (d) | GA 'steibo: 'écraser' > GM 'stivo selon GA 'stupho ⁴ : 'froncer les lèvres' ⁵ |
| | (e) | GMéd. <i>paraḍia'vazo</i> > <i>periḍia'vazo</i> 'se promener' selon <i>περί</i> 'autour' |
| | (f) | GM <i>spaḥan'thi</i> > <i>spaḥan'thi</i> 'groupe d'arbustes' selon IT <i>spade</i> < LT <i>spatha</i> < GA 'spathe' ⁶ : |

L'étymologie populaire peut avoir une dimension diachronique, puisqu'elle conduit à des formes nouvelles, lorsque la forme créée par cette procédure acquiert un tel degré de figement que la base du rapprochement se perd, par ex. *a'yioklima* vs **e'yoklima* « chèvrefeuille ». Elle a également une dimension synchronique, puisqu'il s'agit d'un mécanisme dynamique qui produit des formes au niveau synchronique, par ex. à partir du terme anglais *collateral*⁷ la forme *ko'ladera* est attestée, se basant sur *'kolos* « fesses » et *'adero* « intestin », des contenus qui sont donc étroitement liés.

Il s'agit d'un processus universel, puisqu'il apparaît en une multitude de langues (pour des exemples de langues variées cf. Bauer 2005, Lieberman 2009, 45-55 mais aussi Bauer 2006 / Wikipedia: « folk etymology ») en même temps qu'un processus historique, étant donné qu'il se manifeste à plusieurs périodes d'une langue, ex. GH *ḍiplo'karios* « soldat bénéficiant d'un double salaire » < IT. *dupli-caris* selon *ḍi'plos* « double ». Il s'agit d'un mécanisme de changement analogique (cf. Mc Mahon 1994 ; Fertig 2013) à double statut, puisque tantôt il modifie le mot sans changer son référent, tantôt il fournit une possibilité dénominate, s'assimilant ainsi à un processus de création néologique. Si l'on cherche à déterminer dans quelle mesure l'étymologie populaire consiste en une procédure de formation de nouveaux mots, le constat ci-dessus appelle à reconnaître l'existence d'un continuum dont les deux pôles opposés sont occupés l'un par le renforcement sémantique faible et l'autre par la fonction de dénomination (cf. aussi section « lexicalisation⁸ »).

Les processus liés à l'étymologie populaire interviennent à plusieurs niveaux, puisque les changements qu'ils induisent sont d'ordre phonologique, morphologique, sémantique ou, le plus souvent, résultent de la synergie entre ces trois

1. Prononciation en GM [ana'veno].

2. Prononciation en GM [γavri'o].

3. Prononciation en GM [ḍoli'xos].

4. Prononciation en GM ['stifo].

5. Il s'agit ici d'une erreur étymologique.

6. Prononciation en GM [spa'thi].

7. Dans le vocabulaire technique des contrats bancaires.

8. MÜLLER *et al.* (2015, 1007) l'inclut dans les processus périphériques de formation de mots.

niveaux. Quant à la hiérarchisation des niveaux de fonctionnement et la directionnalité (forme vers sens ou sens vers forme), nous considérons que, probablement, les processus d'assimilation phonologique et de connexion sémantique sont plus ou moins simultanés et non pas séquentiels. Ceci pour la raison que l'assimilation phonologique des différents ingrédients sans parenté au niveau du contenu n'aurait aucun sens, et, inversement, la connexion sémantique sans contrepartie phonologique n'aurait pas beaucoup de succès au sein de la communauté linguistique. Dans la mesure où le départ de l'intervention est souvent spontané et non conscient (cf. section « classement et implications pour le lexique »), il est difficile de prouver l'existence d'un modèle directionnel de dérivation.

Il faudrait préciser à ce point que le changement formel implique la connexion à un autre morphème, entraînant habituellement la modification de la forme de surface de l'unité lexicale, sans que ceci soit toujours nécessaire¹. Par conséquent, en fonction du degré d'action, on peut distinguer entre une formation transparente, lorsque la modification concerne la forme phonologique et/ou la structure morphologique et les ingrédients du mot, ex. *poltrona* > *poli'θrona* « fauteuil », ou uniquement l'orthographe, ex. *κτήριο* (*'ktirio*) > *κτίριο* (*'ktirio*) « bâtiment » ; une formation opaque, lorsque la forme de surface du mot se trouve inchangée, ex. GA 'zeido:ros « fertile » (*zei'a* « blé » + *'do:ron* « cadeau ») > GM 'zidoros « vivifiant » (*zo'i* « vie » + *'doro* « cadeau »).

Le périmètre d'action de l'étymologie populaire est très large, puisqu'il peut s'étendre d'une partie du mot jusqu'au niveau du texte. Plus précisément, elle peut intervenir au niveau du thème des mots simples, GM *fθi'noporo* « automne » > *xi'noporo* ou de dérivés, ex. GMéd. *pate'liða* > GM *peta'liða* « patelle », premiers constituants, ex. GM *filoso'fia* « philosophie » > *ksiloso'fia* « philosophie par raillerie », ou seconds constituants, ex. GH *oma'toklaðon* > GM *ma'toklaðo* « cil ». De surcroît, l'étymologie populaire n'affecte pas uniquement des unités monolexicales mais aussi des composés polylexicaux, ex. fr. *table d'hôte* > GM *tabl'dot* > *'dabl'dot* (rapprochement vers l'angl. *double dot*), des phrases lexicalisées, comme par ex. des titres, des noms de marque déposée, ex. fr. *La vache qui rit* > *la'vas ti'ri* « la vache-fromage », et des collocations, des phraséologismes, par ex. *'frudes el'pides* « espoirs non-fondés » > *'fludes el'pides*, avec rapprochement

1. L'étymologie populaire doit être distinguée: a) des faux amis (cf. ANASTASSIADIS-SYMÉONIDIS & VLETSI 2009), b) des processus liés à la néologie, tels que la construction de mots-valises, où les mots fusionnent selon une opacité morphologique croissante, et des cas de connexion à des nouveaux morphèmes, sans qu'il y ait de changement sur la forme d'un mot morphologiquement complexe (e.g. KOUTITA & FLIATOURAS 2001 ; KARANTZOLA & FLIATOURAS 2004 ; KATSOUDA & NAKAS 2013), c) de la surcorrection, c'est-à-dire une mise en conformité erronée selon des éléments idéalisés de la langue (e.g. KARANTZOLA & FLIATOURAS 2004 ; FLIATOURAS 2018) et de la ré-analyse, qui modifie les frontières morphémiques d'un mot (e.g. ARONOFF 1974 ; HASPELMATH 1994 ; FERTIG 2013 et pour le grec, FLIATOURAS, 2018). Il est à noter que la construction de mots-valises, la connexion à des nouveaux morphèmes, la surcorrection et la réanalyse fonctionnent, sous certaines conditions, comme des mécanismes liés à l'étymologie populaire, dans la mesure où leur intervention contribue à l'association erronée à d'autres morphèmes.

vers *'fluða* « peau du fruit », *'ðu*, *'Kirie*, *fila'kin...* « mets, Seigneur, prison [à ma bouche] » > *'ftu*, *'Kirie*, *fila'ki*, « formule conjuratoire, simulacre de l'action de cracher, Seigneur, prison... », *'eteron e'kateron* « c'est autre chose » > *'eteron ep'apiron* « autre chose à l'infini ».

Classement et implications pour le lexique

Une première distinction des procédés liés à l'étymologie populaire peut s'opérer sur la base de sa cause de départ. Le but de ce processus de nature psycholinguistique consiste à relier une forme inconnue ou peu intelligible à une autre forme, qui est, elle, connue (Anastassiadis-Syméonidis 2003, 56, 59). De cette manière, le locuteur cherche à rendre une forme opaque et irrégulière plus transparente et régulière, avec des procédés qui ne sont certes pas ceux de la linguistique historique ni de l'étymologie savante, mais qui constituent des précieux témoignages sur la structure et l'organisation du savoir lexical des locuteurs.

A) Le départ de ce processus peut être non-conscient, lorsqu'il prend appui sur la difficulté de lier de manière simultanée le signifiant au signifié d'une unité lexicale d'un point de vue phonologique et morphologique. Un nouveau rapprochement est ainsi tenté, où la structure phonologique et morphologique de l'unité existante est associée à un autre élément (ou unité), avec comme objectif un gain au niveau de la transparence. L'implication de processus cognitifs, sur lesquels nous allons revenir plus loin, est ici évidente, et de ce point de vue il s'agirait de processus de nature psycholinguistique. L'étymologie populaire est rangée dans le cadre de ce processus non-conscient, étant donné que les couches populaires ont tendance à ajuster soit les formes savantes caractérisées par un degré de codage réduit, soit les emprunts directs vers des formes plus familières. L'étymologie populaire conduit à des formations qui présentent un niveau élevé de codage, mais aussi à des formations d'un codage réduit. Dans le premier cas, la nouvelle forme acquiert une diffusion importante et finit par se figer, soit dans la norme de la langue, ce qui conduit à la perte de la base du rapprochement, ou encore à un déplacement vers un usage extrêmement savant, ex. *a'yioklima* au lieu de l'initial *e'yoklima* « chèvrefeuille », soit dans un niveau d'usage familier, cas dans lequel il coexiste avec la base du rapprochement qui se maintient dans la norme, ex. *yiuvvar'laki* vs *yiuvare'laki* « boulette de viande et de riz » (cf. Anastassiadis-Syméonidis & Fliatouras 2005).

Le phénomène de changement d'une unité lexicale s'insère dans le deuxième cas, où la forme produite souffre d'un degré de codage et d'acceptabilité réduits, entraînant ainsi une réaction comique de la part de la communauté linguistique, acquérant de cette manière un statut plus sporadique et aléatoire, ex. *'ðu*, *'Kirie*, *fila'kin...* > *'ftu*, *'Kirie*, *fila'ki*, sur la base de la formule conjuratoire *'ftu* « simulacre de l'action de cracher », *ko'biuter* « ordinateur » > *ku'biuter* sur la base du

mot *ku'bi* « bouton », *ó'mi 'yenito* > *ú'mi 'yenito* « que cela n'advienne pas » sur la base du marqueur de négation *ú* « ne pas ».

Comme le fait remarquer Béguelin (2008), le terme « populaire » qui est ici utilisé pour catégoriser la pratique étymologique spontanée des locuteurs n'est pas innocent : il entre en contraste avec « scientifique », et son contenu sociolinguistique se charge de connotations dépréciatives, voire méprisantes sous la plume de certains auteurs. L'étymologie populaire est ainsi opposée à l'étymologie savante, alors que les deux manières d'aborder ce phénomène ne se situent pas, à notre avis, au même niveau : l'apport de l'étymologie savante est lié à sa démarche documentée, exhaustive et axiomatique, alors que l'intérêt principal de l'étude de l'étymologie populaire, en tout cas dans le cadre d'une approche synchronique, se situerait plutôt sur la représentation et construction lexicales, ainsi que sur la manière dont les différents composants du phénomène (phonologique, morphologique, sémantique) interagissent afin d'opérer un rapprochement entre les unités concernées, rapprochement qui sera entériné par la communauté linguistique.

B) Lorsque le départ du processus est conscient, nous pouvons distinguer deux cas :

- a. celui de l'étymologie populaire ludique, lorsque l'unité lexicale change en mot porte-manteau ou acquiert une étymologie nouvelle (cf. Anastassiadis-Syméonidis 1986, 106 ; Katsouda & Nakas 2013), afin de créer des formes non codées ayant un effet comique ou ironique, ex. *ipo-loi'stis* « aide-comptable » sur le modèle de *ipo-loxa'γos* ([sous] [capitaine]) « lieutenant ».
- b. celui de l'étymologie populaire normative, lorsqu'elle émane de positions linguistiques personnelles ou de la politique linguistique étatique et retient comme but le « nettoyage » des éléments étrangers ou l'affirmation de la continuité historique et linguistique, voire la propagande politico-nationaliste. En grec, par ex., nous pouvons attribuer à un effet de rétro-passéisme¹ les efforts — erronés — de rapprochement de mots d'emprunt avec des racines grecques ou les efforts pour rendre savantes certaines formes. Nous attribuons aux effets d'une politique linguistique l'hellénisation des toponymes étrangers (arvanites, slaves et turcs), ex. turc *şahin* « faucon » > *E'xinos* (selon GA *e'khinos*, GM *axi'nos* « oursin ») (Syméonidis 2010). Sur la base de ces éléments, nous pouvons observer que les termes « étymologie populaire » et « fausse étymologie savante » devraient être considérés comme des hyponymes, avec un terme plus général, celui de « parétymologie », qui constituerait l'hyperonyme.
- c. Un autre élément qui va dans le sens de ce classement, vient du fait que, dans le cadre d'une approche synchronique et du point de vue du locuteur natif et de la communauté linguistique, nous pouvons aisément constater que le

1. Courant de pensée prônant le retour à l'antiquité et qui frappe de manière exacerbée certains locuteurs, selon lesquels tout prend sa source au grec ancien. Il peut être traduit par le terme « archéolâtre » même si ce mot est plutôt un néologisme en français, ou encore « atteint d'archéolâtrie aiguë ».

résultat d'un rapprochement lié à l'étymologie populaire s'intègre parfaitement dans la connaissance lexicale du locuteur, avec le même statut que les mots n'ayant pas fait l'objet de ce type de rapprochement, et qui sont le résultat d'une « vraie » relation étymologique. Afin de mieux illustrer ce point, nous pouvons brièvement rapporter l'étude de Anastassiadis-Syméonidis & Voga (2010) auprès de locuteurs natifs du grec, tous d'un niveau d'éducation universitaire, ayant participé à une expérience de jugement de « grécité » de mots d'emprunt. Ces mots d'emprunt présentaient une structure morphologique résultant d'un procédé d'étymologie populaire, par ex. *poli'θrona* « fauteuil », ou *sintri'vani* « fontaine décorative¹ », et nous leur avons attribué le terme pseudo-préfixés, puisqu'ils présentent une structure morphologique de préfixé (*poli-* 'poly-' dans le premier cas et *συν-* (*sin*) 'syn-' dans le second). Les résultats de cette étude montrent que les scores de « grécité » qui leur ont été attribués sont très proches de ceux de mots dont l'origine grecque ne fait aucun doute.

Fait intéressant, ces pseudo-préfixés obtiennent des scores de grécité significativement supérieurs aux pseudo-suffixés (ex. *me'talio* « médaille », *mi'dralio* « mitraille »). Si, en suivant la dichotomie traditionnelle entre étymologie populaire et savante, on abordait ce type de résultat uniquement au travers de l'erreur étymologique de la part du locuteur, ou, alternativement, uniquement à travers la raison de départ, ici, la volonté d'un pouvoir socio-politique d'attribuer davantage de « grécité » à des mots d'origine étrangère, cela ne suffirait pas à saisir le phénomène dans sa complexité. Il nous semble beaucoup plus pertinent d'attribuer à l'hyperonyme « parétymologie », un caractère systématique, présentant une certaine forme de régularité, dans le sens de l'analogie, non seulement au niveau phonologique mais aussi morphosémantique, et ceci indépendamment de la raison de départ, consciente et voulue, résultat d'une intervention verticale, ou pas. La section suivante fournit quelques éléments en faveur du caractère systématique de ce type de fonctionnement.

Le fonctionnement systématique de l'étymologie populaire

L'étymologie populaire est souvent vue comme une sous-catégorie d'analogie aléatoire qui affecte le niveau phonologique (cf. entre autres, McMahon 1994). Il y a néanmoins des indices en GM qu'elle possède un certain degré de systématisme et de régularité, non seulement au niveau phonologique mais aussi morpho-sémantique, liées à la base du rapprochement et la forme d'intervention phonologique et morphologique.

1. La fontaine « appareil où l'on peut boire de l'eau » est désignée en grec par une autre unité lexicale (*psiktis*).

Caractéristiques de la base des mots faisant l'objet d'un processus d'étymologie populaire

Certaines catégories de mots caractérisées par un degré élevé d'opacité morpho-sémantique constituent un terrain privilégié pour l'étymologie populaire, comme (a) les mots d'emprunt, (b) les formes savantes sous forme de mots ou phraséologismes et (c) les mots difficiles à prononcer (cf. aussi Anastassiadis-Syméonidis 1994, 55, 109 ; Anastassiadis-Syméonidis 2003a, 56-59 ; Karantzola & Fliatouras 2004¹):

(3)	(a)	FR	<i>tire-bouchon</i>	>	GM tirbu'son > tripi'son 'tire-bouchon' selon tri'po 'trouer'
		IT	<i>grissino</i>	>	GM kri'tsini 'gressin' selon l'onomatopée 'kritis 'scrunch'
		TC	<i>yuvarlak</i>	>	GM yiuware'laki 'boulette de viande et de riz' selon va'reli 'tonneau'
(b)		GA	ai'goklima ²	>	GMéd. a'yioklima 'chèvrefeuille' selon 'ayios 'saint'
		GH	pareyi'aliós	>	GMéd. peri'yiali 'rivage' selon peri- 'autour' et yia'los 'rivage'
		GA	fði'nopo:ron ³	>	GM xi'noporo 'automne' selon 'xino 'verser, partir en vrille'
(c)		IT	<i>poltrona</i>	>	GM poli'θrona 'fauteuil' selon poli- 'poly-' et 'θronos 'trône'
		GA	andrakhle ⁴ :	>	GM a'drakla 'herbe aux vertus aphrodisiaques' selon 'adras 'homme' et -akla (suffixe augmentatif)
		PS	<i>kārānfel</i>	>	GH kari'ofilon 'girofle' selon 'karion 'noix' et 'filon 'feuille'

Enfin, nous pouvons aussi distinguer des sous-catégories des exemples présentés en (3), marquées par une tendance vers le procédé d'étymologie populaire :

- les toponymes d'emprunt, ex. *Clair Mont* > *Xle'mutsi*⁵
- des éléments savants de la tradition ecclésiastique, comme les surnoms de la Vierge, ex. *Pana'γia i Katapolia'ni* « la Vierge qui court les villes » > *Pana'γia i Ekatoapilia'ni* « la Vierge aux cent portes », *Pana'γia i Goryoe'pikoos* « la Vierge qui écoute vite » > *Pana'γia i Goryoi'pikoos* « la Vierge qui obéit vite ».
- les chansons étrangères, ex. *babo'leo* > *babo'leiro*.

Niveaux de réalisation

L'étymologie populaire est un processus qui se réalise simultanément à différents niveaux d'analyse linguistique (phonologie, morphologie et sémantique), basée sur

1. Dans certains cas les unités lexicales peuvent appartenir dans deux catégories, ex. les UL *poltrona* et *kārānfel*, présentés en (c), appartiennent à la fois dans la catégorie des emprunts et dans celle des mots difficiles à prononcer.

2. Prononciation en GM [e'γoklima].

3. Prononciation en GM [fði'noporo].

4. Prononciation en GM [an'draxli].

5. Un grand nombre d'exemples se trouvent dans SYMÉONIDIS (2010).

un continuum d'action qui s'étend par niveau et qui conduit à une intervention variant entre le degré zéro jusqu'à une intervention multiple. La manière dont se réalise cette intervention est souvent systématique, créant ainsi des régularités, ce qui induit un fonctionnement relativement prévisible dans certaines catégories. Dans cette section, nous examinons la manière dont l'étymologie populaire fonctionne dans les différents niveaux.

Étymologie populaire phonologique

Dans la littérature internationale, l'étymologie populaire est le plus souvent considérée comme un processus phonologique analogique (même s'il existe des exceptions, ex. la morphologie naturelle de Kilani-Schoch et Dressler 2005). En effet, son action phonologique est considérable, puisque la nouvelle forme rappelle de manière automatique le mot-cible, sur le plan morphologique et sémantique. Ce rappel est total lorsqu'il se base sur la fonction paradigmatic et induit comme effet la substitution de morphème, ex. GH *to'ksarion* > GMéd. *do'ksa-ri(on)* « archet », ou partiel, lorsqu'il se base sur la construction de mots-valises, ex. GMéd. *pediklo* > GM *burdu'klono* « s'emmêler les pinces, brouiller un récit » selon *ber'devo* « confondre » et *pe(r)di'klono*¹ « faire un croche-pied ». Dans ce qui suit, nous présentons quelques arguments en faveur de son caractère systématique.

Degré d'action

Si nous classifions l'action phonologique dans une échelle ascendante, nous repérons les points suivants :

- Degré zéro : l'étymologie populaire s'apparente à un processus non perceptible d'identification de segments homonymes, sans qu'il y ait de changements phonologiques hormis ceux qui sont prévus. Par ex. dans le mot (GH) *μαμμόθρεπτος* « élevé par la grand-mère », excepté la dissimilation prévisible /pt/ > /ft/, le premier constituant a été rapproché au mot *μαμά* « maman » et a été transformé en GM en *μαμόθρεπτος* « fils à maman », vu la similitude des bases dans les mots *μάμμη* « grand-mère » et *μαμά* « maman ».

Quelques autres exemples :

- | | | |
|-----|-----|---|
| (4) | (a) | GH <i>'onayros</i> ('onos 'âne' + a'γros 'champ') > GM <i>'onayros</i> ('onos 'âne' + 'ayrios 'sauvage') |
| | (b) | VN <i>paraman</i> > GM <i>para'mana</i> 'épingle à nourrice' (para 'para' + 'mana 'mère', cf. et homonyme para mana 'nourrice') |
| | (c) | GA <i>'zeido:ros</i> 'fertile' (zei'a 'blé' + 'do:ron 'cadeau') > GM <i>'zidoros</i> 'vivifiant' (zo'i 'vie' + 'doro 'cadeau') |
| | (d) | GH <i>oma'toklaδon</i> ('oma 'œil' + ki'laδes 'la partie sous les yeux' > GM <i>ma'toklaδo</i> ('mati 'œil' + kla'di 'branche') |

1. Koutita & Fliatouras (2001) désignent ces cas avec le terme « pseudo-mots-valises ».

• Phonétique: il s'agit ici d'une adjonction, suppression ou substitution de son à l'initiale ou à l'intérieur du mot. Ci-dessous, les différents cas possibles avec des exemples¹:

(5)	(a)	Adjonction à l'initiale:	GA 'ozos 'noeud' > GM 'rozoz selon πίζα 'racine' ou 'roz 'rose'
	(b)	Adjonction à l'intérieur:	GMéd. peđi'klo > GM perđi'klono 'faire un croche-pied' selon ber'devo 'confondre'
	(c)	Substitution à la syllabe initiale:	GH to'ksarion > GMéd. do'ksarion 'archet' selon 'đoksa 'gloire'
	(d)	Substitution médiane:	IT <i>medaglia</i> > GM me'talio 'médaille' selon GA 'metallon 'métal' (GM 'metalo)
	(e)	Suppression à la syllabe initiale:	GH ipo'deno > GM po'deno 'chausser' selon 'pođi 'pied'
	(f)	Suppression médiane:	GA oike:'te:rion ² 'habitation' > GMéd. 'ktirion 'bâtiment' selon 'ktizo 'bâtir'

• Syllabique: elle consiste en une suppression, remplacement ou permutation d'une syllabe, à l'initiale ou à l'intérieur du mot ; ci-dessous quelques exemples caractéristiques:

(6)	(a)	Substitution à la syllabe initiale:	VN <i>stramazada</i> > GM stroma'tsađa 'couchage par terre' selon 'stroma 'matelas'
	(b)	Substitution médiane:	GA Nere:'ida 'Néréide' ³ (acc.) > GMéd./GM ne'raiđa 'fée' selon ne'r'o 'eau'
	(c)	Suppression médiane:	ALL <i>papperlapap</i> (onomatopée) > GM parla'pipa selon 'parla 'bagou' et peut-être 'pipa 'pipe'
	(d)	Permutation à la syllabe initiale:	GMed. pate'liđa (acc.) de 'patella 'assiette' > GM peta'liđa 'patelle' selon 'petalo 'fer à cheval'
	(e)	Permutation médiane:	GMéd. kami'lafki > GM kali'mafxi 'chapeau des popes orthodoxes' selon 'kalima 'couverture, protection' et peut-être af'xenas 'nuque'

• Plurimorphémique: le processus lié à l'étymologie populaire consiste à créer une structure complexe de pseudo-préfixé (ex. IT *poltrona* > poli'throna « fauteuil »), pseudo-suffixé (an'đraxli > a'drakla « herbe avec des vertus aphrodisiaques ») ou pseudo-composé (fđi'noporo > xi'noporo « automne »).

1. L'application de plusieurs règles phonologiques n'est pas exclue, ex. GA oike:'te:rion « habitation » > GM 'ktirio « bâtiment » (f) où il y a suppression à l'initiale et à l'intérieur du mot.

2. Prononciation en GM [iki'tirion].

3. Dans la mythologie, les Néréides sont des nymphes marines, filles de Nérée et de Doris.

Environnement de l'action de l'étymologie populaire

L'étymologie populaire prend effet dans des environnements phonologiques concrets, selon des règles de fonctionnement que l'on pourrait comparer aux règles phonologiques.

- Les emprunts et les éléments savants contenant des groupes consonantiques périphériques notamment à l'intérieur du mot, provenant d'autres langues ou de phases antérieures du grec, où les contraintes phonotactiques sont différentes. Une telle catégorie serait constituée de manière majoritaire par les mots qui présentent des fricatives. Même si, dans un tel cas, des processus phonologiques curatifs auraient pu intervenir, comme la simplification et l'épenthèse, l'intervention de l'étymologie populaire semble fournir un gain supplémentaire, vu qu'elle rend le mot plus transparent sémantiquement et morphologiquement et permet la reconnaissance de morphèmes familiers. Par ex., le mot GA *an'drakhle*: « herbe avec des vertus aphrodisiaques », d'étymologie incertaine, portant un groupe consonantique à trois membres (dans d'autres cas à deux) a été relié au mot *'andras* « homme ». Quelques exemples de ce type ci-dessous :

- | | | |
|-----|-----|--|
| (7) | (a) | FR <i>blague</i> > GM 'plaka 'blague' selon 'plaka 'disque de grammophone' |
| | (b) | GA phthi'nopo:ron ¹ > GM xi'noporo selon 'xino 'verser, partir en vrille' |
| | (c) | IT <i>grissino</i> > GM kri'tsini 'gressin' selon l'onomatopée 'krits 'scrunch' |
| | (d) | TC <i>yağma</i> 'butin' > GMéd. <i>ḡaiyu'mas</i> 'pilleur' |
| | (e) | IT <i>poltrona</i> > GM poli'throna 'fauteuil' (poli- 'poly-' + 'thronos 'trône') |
| | (f) | PS <i>kārānfel</i> > GH kari'ofilon 'girofle' ('karion 'noix' + 'filon 'feuille') |

- L'homonymie et la paronymie des morphèmes conduit souvent à des changements morphologiques dans la langue (Anastassiadis-Syméonidis 2003a, 58). Par ex. le mot *ma'nes* provenant du mot turc *man(i)* « entrave, obstacle — genre de musique populaire » et le mot *a'man* « zut, plainte » sont des paronymes, ce qui explique l'adjonction du *a-* à *ma'nes*, pour donner le nom *ama'nes* « chanson orientale répétitive et plaintive », créé à partir de ce procédé d'étymologie populaire vers le mot *a'man*. Aussi, le mot d'étymologie incertaine (GA) *arseni'ko* « arsenic » est assimilé au nom *arseni'ko* « mâle », sur la base du lien associatif métaphorique entre la puissance de l'homme et de ce poison. Enfin, dans le mot (GA) *mu'riki* « arbre présent en front de mer », nous observons une homonymie avec le suffixe *-'iki* et une paronymie avec le mot *'armi* « liquide salé destiné à la conservation des aliments » conduisant à une ré-analyse du mot et la nouvelle forme *armi'r-iki* « tamaris, arbre présent en front de mer ».

1. Prononciation en GM [fθi'noporo].

Règles phonologiques

L'intervention de l'étymologie populaire sur la forme phonologique des morphèmes ressemble souvent à celle de règles phonologiques responsables des changements de forme, à la nuance près que le but de celle-ci n'est pas de réparer mais d'assimiler la base au mot-cible. De telles règles, relatives aux voyelles (a) et aux consonnes (b), sont présentées ci-dessous :

(8) (a)	Fermeture :	IT <i>clavicembalo</i> 'clavecin' > GMéd. klaví'tsibalon 'clavecin' selon tsi'bo 'piquer'
	Ouverture :	TC <i>kurak</i> > GM kora'k-iazo 'avoir la pépie' selon ko'raki 'corbeau'
	Arrondissement :	GM 'refulas 'tornade' > 'rufulas 'siroteur' selon ru'fo 'siroter, aspirer'
	Effacement :	GH oma'devo > GMéd. ma'zevo 'rassembler' selon ma'zi 'avec, ensemble'
	Simplification :	GH ayri'abelos > GM a'γrabeli 'vigne vierge' selon a'γros 'champ'
	Syncope :	GA oike:'te:rion ¹ 'habitation' > GMéd. 'ktirion 'bâtiment' selon 'ktizo 'bâtir'
	Assimilation :	GMéd. me'rono 'apprivoiser' > mo'rono 'devenir idiot' selon mo'ros 'idiot'
(b)	Assourdissement :	IT <i>grissino</i> > GM kri'tsini 'gressin' selon l'onomatopée 'krits /'krits/
	Sonorisation	'scrunch' IT <i>battitura</i> > GM pati'tura 'décalcomanie' selon pa'to 'presser' IT <i>medaglia</i> > GM me'talio 'médaille' selon 'metalo 'métal' GMéd. 'sfika > GM 'sfiga 'guêpe' selon 'sfigo 'serrer, contracter' ou 'Sfiga 'Sphinx'
	Fricativisation :	FR <i>mitraille</i> > mi'ðralion 'mitraille' selon 'miðros 'obus'
	Affrication :	GA em'pis > GM tsi'biða 'pince' selon tsi'bo 'piquer'
	Assimilation :	VN <i>turlulu</i> 'simple d'esprit' > GM surlu'lu 'écervelée' selon 'surno 'tirer, trainer' et -u
	Dissimilation :	GH esxa'toyiros 'homme très avancé en âge' > GM ska'toyeros 'sale vieux' selon ska'to- 'merde'
	Permutation :	GMéd. pate'liða (acc.) > GM peta'liða 'patelle' selon 'petalo 'fer à cheval' GMéd. kami'lafki > GM kali'mafxi 'chapeau des papes orthodoxes' selon 'kalima 'couverture, protection'

1. Prononciation en GM [iki'tirion].

Fonctionnement du lexique mental: du perceptuel au conceptuel

Concernant les données présentées dans les sections précédentes décrivant l'intervention de l'étymologie populaire sur la forme phonologique des mots, nous pouvons remarquer que, pour tous les cas décrits (intervention phonétique, syllabique, polymorphémique¹), nous observons que le gain de transparence se réalise grâce à l'appariement à un nom (parfois un verbe ou adverbe) qui correspond à un référent connu par le locuteur, par ex. *'stroma* « matelas » pour *stroma'tsaða* (6a), *ne'ro* « eau » pour *ne'raiða* « fée » (6b), *'podì* « pied » pour *po'deno* « chausser » (5e), ou *'ktizo* « bâtir » pour *'ktirio* « bâtiment » (8a). Ce type de fonctionnement est, comme le rappelle Béguelin (2002), étroitement lié à l'antanaclase, la paronomase, aux phénomènes de rime ou d'allitération, qui constituent des moyens éprouvés pour révéler ou pour suggérer, même fugitivement, une parenté sémantique ou étymologique entre les expressions concernées². L'efficacité de ce procédé repose sur un principe symbolique selon lequel l'affinité des signifiants serait de nature à cautionner l'affinité des signifiés, et à soutenir une assertion d'identité ou de ressemblance entre les référents évoqués (Béguelin 2002).

Concernant les preuves expérimentales de ce type de fonctionnement chez le sujet parlant, il existe de nombreuses études qui démontrent que le rapprochement entre deux formes qui ont un signifiant similaire mais aucun lien morphologique, historique ou étymologique, non seulement présente une réalité psychologique, chose à laquelle on peut logiquement s'attendre, mais aussi produit des effets très difficiles à distinguer des effets réellement morphologiques, en tout cas pour ce qui concerne les premiers stades du traitement langagier. Nous faisons référence ici à l'effet de pseudo-suffixation, obtenu dans des protocoles expérimentaux de chronométrie mentale, notamment en amorçage masqué auprès de locuteurs natifs. Cette relation de pseudo-suffixation concernerait, dans ce type de protocole (ex. Rastle, Davis & New 2004) des UL du type *corner-corn* « angle — maïs » (pour la version française, Longtin & Meunier 2005 : *baguette-bague*). Si l'un des auteurs de ces lignes a consacré plusieurs publications sur la nécessité et la possibilité, grâce à la planification expérimentale adéquate, de distinguer entre les effets réellement morphologiques et les effets pseudo-morphologiques (pour une revue sur la question et les données, cf. Voga & Giraudo 2017), force est de constater que l'effet de pseudo-suffixation (ex. *corner-corn*) et ses variantes, démontre que les processus cognitifs impliqués dans la pseudo-suffixation partagent au moins un niveau de traitement avec ceux responsables du traitement des mots construits. Ceci est vrai, en particulier dans une perspective *bottom-up*, pour des processus perceptifs dont le but est de capter la similarité et/ou saillance perceptuelle entre deux stimuli langagiers. La principale conclusion de ces études (ex. Rastle *et al.* 2004 ; Longtin

1. À l'exception peut-être de l'action au degré zéro, où la similitude formelle est déjà suffisante pour justifier le rapprochement qui s'opère grâce à l'étymologie populaire.

2. C'est le cas dans le célèbre dicton latin *nomen omen* (« le nom est un présage »).

& Meunier 2005) amène à postuler que le mécanisme qui sous-tend le traitement en ligne des UL morphologiquement complexes est la décomposition, c'est-à-dire une segmentation en morphèmes « obligatoire » (Rastle & Davis 2008 ; Rastle *et al.*, 2004 entre autres). Pour ces auteurs, « la décomposition morphologique est un processus qui est appliqué à tous les stimuli qui sont morphologiquement structurés, indépendamment de leur caractéristiques lexicaux, sémantiques ou syntaxiques » (Rastle & Davis 2008, 949). Ce fonctionnement serait de nature proprement morphologique et devrait être pris comme tel.

Cette position est formellement contestée par d'autres chercheurs (Voga & Giraudo 2017 ; Giraudo & Voga 2016 ; pour des données provenant du grec, cf. Voga, Giraudo & Anastassiadis-Syméonidis 2012). Notre but ici n'est pas de poursuivre cette discussion, mais plutôt de relever que le rapprochement entre UL n'ayant pas de rapport sur le plan étymologique ou sémantique mais qui sont similaires sur le plan du signifiant, et *a fortiori* entre celles qui créent l'impression d'une parenté morphologique, ne serait-ce qu'en apparence, correspond à des processus cognitifs bien identifiés. Ces effets de pseudo-préfixation relèvent majoritairement de processus perceptifs, même s'ils sont souvent interprétés comme étant des effets proprement morphologiques (ex. Amenta & Crepaldi 2012 ; Rastle *et al.* 2004). Cette interprétation se base sur l'hypothèse sous-jacente suivante : si des effets significatifs d'amorçage peuvent émerger à partir de la structure de surface des mots (de leur forme) exclusivement, que ces mots soient morphologiquement complexes ou pas, alors, la morphologie n'est pas codée dans le lexique mental, mais dans ces routes d'accès (Giraudo & Dal Maso 2017, 3). Dans cette littérature pléthorique¹ concernant le traitement en ligne des mots complexes, il existe tout de même des résultats expérimentaux mettant en évidence des différences qualitatives entre les deux types d'effets, par ex. l'étude de Giraudo & Voga (2016) sur les thèmes liés du français, ex. *virus-viral* (cf. aussi Giraudo & Dal Maso 2016 ; Voga & Giraudo 2017).

Pour récapituler, le type d'effet qui nous intéresse ici, l'effet de pseudo-préfixation, est capté par le premier niveau de traitement d'UL, responsable de la reconnaissance de la régularité et saillance perceptives de l'input. Le type d'unités captées par ce premier niveau serait plus proche des « morphomes » dans le sens d'Aronoff (1994), c'est-à-dire des unités d'accès qui n'appartiennent pas au niveau morphologique. Le fait que ce type de similarité perceptive soit repérable par le système de traitement langagier et qu'il facilite l'accès aux niveaux plus centraux, où le sens des UL est représenté, suffit pour décrire, en termes fonctionnels, le gain de traitement que représentent les rapprochements effectués à l'aide de l'étymologie populaire.

1. Pléthorique certes, mais qui semble parfois faire des choix incompréhensibles d'un point de vue linguistique, notamment en feignant d'ignorer la distinction entre les bas niveaux de traitement en ligne (perception de lettres, sons, etc.) et les niveaux plus centraux, ex. sémantiques et conceptuels, où le sens des UL est représenté.

Cependant, le fait qu'il existe des différences de traitement et de représentation entre les mots morphologiquement complexes et pseudo-complexes ne signifie pas qu'il n'existerait aucun lien entre les niveaux perceptifs du traitement et les niveaux conceptuels. Des travaux récents démontrent qu'un rapprochement entre des formes n'ayant pas de lien morphologique ou étymologique est tout à fait possible, grâce au mécanisme de combinaison conceptuelle (Gagné & Spalding 2014b ; 2015, voir aussi Smith & Osherson 1984). Ce type de mécanisme se fonde sur un processus de construction du sens et correspond à une tentative d'identifier la relation sur laquelle se base le lien entre les constituants, y compris pour des (nouveaux) mots composés opaques, où, par définition, le sens du composé ne peut pas être le résultat d'un processus de construction du sens (Gagné & Spalding, 2014a). Ce type de fonctionnement témoigne de la grande plasticité et adaptabilité du lexique mental, capable d'attribuer du sens à des stimuli qui en sont dépourvus. C'est précisément en ceci que les processus liés à l'étymologie populaire comportent une dimension proprement morphologique, comme nous le verrons ci-dessous.

Étymologie populaire et morphologie

La bibliographie internationale attribue à l'étymologie populaire une dimension phonologique et, indirectement, morphologique, dans la mesure où la forme phonologique des morphèmes est modifiée, de manière à ce qu'ils puissent être reliés à de nouveaux morphèmes. Il y a néanmoins des indices, pour le GM, que ce type de fonctionnement possède une dimension proprement morphologique. Plus précisément, si nous classifions l'intervention morphologique à l'aide d'une échelle ascendante, nous pouvons repérer quatre points distincts : substitution, changement de genre, restructuration et lexicalisation. Nous allons examiner dans le détail ces quatre niveaux d'action morphologique.

Substitution

Des mots morphologiquement simples font l'objet d'un rapprochement sans changement au niveau du contenu, résultant à la naissance d'un nouveau pseudo-allomorphe thématique, ex. *erpe'to* « reptile » > *serpe'to* « reptile » selon le v. *sernome* « ramper ». Par ailleurs, des mots affixés ou des composés sont ré-analysés sur la base de constituants plus familiers et productifs, qui viennent en remplacer d'autres, plus savants, d'habitude sans changement au niveau du contenu, ex. *filoso'fia* « philosophie » > *ksiloso'fia* « philosophie (par raillerie) ». Ce phénomène rappelle fortement celui de la substitution morphologique, ex. *oðodal'ya* « dentalgie » vs *po'noðodos* « mal de dents », à la différence près que les segments ne sont pas substitués mais font néanmoins l'objet d'un rapprochement. Nous pouvons distinguer les cas suivants :

- Substitution des constituants du mot : en GH, le mot *esxa'toyiros* qui dénote l'homme d'un âge avancé a *ésxatos* « dernier » comme premier constituant et *ýeros* « vieux », comme second. Outre le remplacement, prévisible, de l'allomorphe savant *yir-* par *ýeros* « vieux », on peut observer que le premier constituant (*ésxatos* « dernier ») est remplacé par le constituant *ska'to* « merde », sur la base du sentiment désagréable que les personnes âgées peuvent provoquer à cause de leur comportement difficile.
- Substitution de prépositions/préfixes et préfixoïdes : la substitution de prépositions par d'autres, plus productives, moins savantes et qui apportent un bénéfice prédictif de la fonction sémantique, ne serait-ce qu'indirectement, est systématique dans les mots du GA et du GH. Par ex., les préfixes *hupo-* 'hypo-' et *epi-* 'épi-' dans *hu'pe:nemos*¹ « abrité du vent », *epithu'mo*² : « désirer », *epakro'o:mai*³ « entendre avec beaucoup d'attention » se rattachent en GM ou en GH selon le préfixe *apo-*⁴ et ces unités lexicales changent en *απάνεμος* « abrité du vent », *apoθi'mo* « souffrir de l'absence de quelqu'un, s'ennuyer de quelqu'un » et *afu'grazome* « tendre l'oreille ». Respectivement, GH *paryi'ali* (adj.) « au bord de la mer » change en GM *peri'yiali* (subst.) « bord de mer », GH *eksomolo'yo* « confesser » en GM *ksomolo'yo* « confesser » (*eks-* > *ekso-*). Enfin, la substitution du segment *meso-* « au milieu » par *miso-* « semi- », ex. GM *meso'fori* > *miso'fori* « jupon », GM *meso'chori* > *miso'chori* « village entre deux villages », GM *meso'ximono* > *miso'ximono* « milieu de l'hiver ».

Attribution de genre et de catégorie grammaticale

Le genre et/ou la catégorie grammaticale d'un mot sont souvent attribués sur la base du rapprochement avec un autre mot :

-
- (9) IT *banco* > GM *'banka* 'banque au jeu de cartes' selon *'banka* 'banque'⁵
 NL *peronospora* > GM *pero'nosporos* 'mildiou' selon *'sporos* 'grain'
 IT adj. *fasulo* 'faux' > GM nom *fa'suli* 'sujet inattendu' selon *fa'suli* 'haricot'
 GA adj. *pa'no:le:s*⁶ 'destructeur' > GM nom *pa'noli* 'peste' selon *pa'nukla* 'peste'
-

1. Prononciation en GM [i'pinemos].

2. Prononciation en GM [epiθi'mo].

3. Prononciation en GM [epakro'ome].

4. Nous rappelons que les préfixes du GM viennent de prépositions ou d'adverbes du GA (Fliatouras 2018).

5. Concernant l'attribution de genre aux mots d'emprunt, cf. ANASTASSIADIS-SYMÉONIDIS (1994) et CHRISTOFIDOU (2003).

6. Prononciation en GM [pa'nolis].

Restructuration

Nous observons souvent une intervention au niveau de la structure interne des mots conduisant à une modification de cette structure. Les différents cas de restructuration liés à l'étymologie populaire peuvent être représentés sur une échelle ascendante de la manière suivante :

- Le début ou la fin du mot est assimilé à un affixe dérivationnel, ce qui résulte en la construction de mots où seul l'affixe est reconnaissable (pas la base). Ce phénomène rappelle les paradoxes structuraux de Pesetsky (1979¹), par ex. lorsque le mot italien *allaburlesca* a été interprété comme « type de langue ludique », ce qui a rendu possible l'assimilation du segment *-esca* au suffixe du grec *-'ezika*. En d'autres termes, cela revient à créer *ad hoc* une base de dérivation *alaburn-* sur laquelle le suffixe *-'ezika* peut s'appliquer, de manière à ce que ce mot puisse être intégré dans la classe des langues parlées, par ex. *ki'n-ezika*² « chinois ». Ci-dessous, quelques exemples caractéristiques :

-
- (10) δία- : TC *yağua* 'butin' > GMéd. δία-γυ'mas 'pilleur' > GM δία-γυ'mizo 'piller'
- 'ini : FR *Ripolin* (nom propre) > GM ripo'l-ini 'ripolin'
- 'aki : TC *perdah* > GM ber'daxi > ber'd-aki 'bastonnade'
- 'ula : IT *ranuncolo* > GM nera'g-ula 'renoncule'
- 'u : VN *turlulu* 'simple d'esprit' > GM surlu'l-u³ 'écervelée, trainée'
-

- Des mots composés sont transformés en mots affixés ou l'inverse. L'association entre forme et contenu n'est cependant pas toujours claire, et peut être étendue sur un continuum, dont l'une des deux extrémités est occupée par des exemples comme ceux présentés en (11). Par ex. le composé GMéd. *eksopor'tizo* « sortir de la porte » est ré-analysé en GM comme *ksepor'tizo* « partir », étant donné que le préfixe *kse-* comporte l'instruction sémantique d'éloignement, alors que, par opposition, dans le mot *kse'fteri* « épervier », la contribution sémantique de *kse-* est nettement occultée. Cette diminution de la transparence sémantique est plus prononcée dans les exemples présentés en (11).

-
- (11) GH oksi'pterion (oksi- 'aigu' + pte'ron 'aile' + -ion 'suff.dér.') > GMéd. kse'fteri 'épervier' (kse- + fte'ron 'plume')
- GH eks-omolo'yo > ekso-molo'yo 'confesser' (eks- > ekso- 'dehors') omolo'yo
- GH 'plimira ('plimi 'élévation du niveau de rivières, de mer' + -i'ros) > pli'mira 'inondation' (plin + 'mirome 'couler')
-

1. Selon le modèle de CORBIN (1987), il s'agit d'une catégorie de mots dont la structure est [-simple, -construite], autrement dit, entre les mots simples et les mots construits.

2. Ce phénomène ressemble à la réanalyse postérieure d'internationalismes, ex. fr. *alpinisme* > gr. *alpini'sm-os* > *alpin-i'smos* (cf. ANASTASSIADIS-SYMÉONIDIS & GALANI 1995), à la différence près que la reconnaissance du suffixe n'est pas vraie mais fausse.

3. Cette forme renforce le sentiment que la base est reliée au verbe « trainer ».

• Des mots morphologiquement non construits sont transformés en pseudo-construits¹. Ce type de phénomène est fréquent dans le changement morphologique, par ex. GA *mu'riki*² (mot simple) est ré-analysé en pseudo-dérivé GM *'armi* « liquide salé destiné à la conservation des aliments » + *-'iki*. Cette procédure est similaire à celle de la dérivation régressive, ex. TC *çardak* > GM *tsar'dak-i* « piaule » > *tsar'd-aki* > *tsar'di*. La différence cependant ici réside dans le fait que la dérivation régressive conduit *ad hoc* à un mot nouveau, alors que le procédé d'étymologie populaire crée une nouvelle pseudo-base, ou alors la base existante est reliée à une autre. Nous pouvons distinguer deux cas principaux de restructuration : (a) transformation d'un mot morphologiquement simple en pseudo-dérivé et (b) transformation d'un mot morphologiquement simple en pseudo-composé :

-
- (12) (a) FR *mitraille* > GM *mi'dralio* ('miðros 'obus' + *-'alio*)
 FR *capsule* > GM *ka'psilio* 'capsule' (kaps- / 'kapsa 'capsule' + *-'ilio*)
 IT *grissino* > GM *kri'tsini* 'gressin' (onomatopée 'kritis / 'kritis / 'scrunch' + *-'ini*)
 FR *gamin* > GM *xa'mini* 'gosse négligé' ('xamo 'par terre' + *-'ini*)
 GH *me'morion* > GMéd. *mni'muri* ('mnima 'tombeau' + *-'uri*)
 GA *an'drakhle*³: (étymologie incertaine) 'herbe avec des vertus aphrodisiaques' > GM *a'drakla* ('adras 'homme' + *-'akla*)
 LT *defendere* 'défendre' > GMéd. *ðiafe'ntevo* 'gouverner' (ðia- + *a'fentis* 'maître')
 IT *poltrona* > GM *poli'throna* 'fauteuil' (poli- 'poly-' + *'thronos* 'trône')
-
- (b) PS *kärämfel* > GH *kari'ofilon* 'girofle' ('karion 'noix' + *'filon* 'feuille')
 ALL *papperlapap* (onomatopée) > GM *parla'papa* ('parla 'bagou' + peut-être *'papa* 'pipe')
 TC *kulampara* 'homosexuel' > GM *koloba'ras* 'homosexuel vulg.' ('kolos 'fesses' + *'bara* 'barre')
-

Lexicalisation

L'étymologie populaire peut, plus rarement, fonctionner comme un mécanisme de lexicalisation, étant donné qu'elle transforme des phrases en des composés⁴:

-
- (13) SL *vol* 'boeuf' + *void* 'écorcher' > GM *volo'derno* ('volos 'boule' + *'derno* 'battre') 'battre la campagne'
 GM *'male* 'vrase Tr.lit. 'bouillir', Tr.Fr. 'remue-ménage' > *malio'vras* selon *ma'lia* 'cheveux'
-

En conclusion, sur la base du classement ci-dessus, le degré d'action morphologique de l'étymologie populaire, y compris dans les cas de simple changement

1. Plus rarement, le phénomène inverse est observé, c'est-à-dire que la structure morphologique des mots complexes disparaît suite à l'application des règles phonologiques, ex. GMéd. *ipo'deno* > GM *po'deno* « chausser », GH *po'yon-ion* « menton » (diminutif du GA 'po:go:n « menton ») > GMéd. *pi'guni(n)* « menton ».

2. Prononciation en GM [mi'riki].

3. Prononciation en GM [an'draxli].

4. L'intervention de l'étymologie populaire peut aussi conduire à des formes composées erronées par rapport à la forme de départ, par ex. dans les traductions d'emprunts, ex. IT *olio di ricino* > *retsi'nolado / riti'neleo* (forme savante) « huile de ricin ».

formel que nous avons vus en 3.2.1.1, peut être représenté à l'aide du continuum ci-dessous :

(14)

Niveaux de réalisation			
Substitution	Attr. de genre/cat.	Restructuration	Lexicalisation
morphématique	catégorielle (changement de nombre, genre)	structurale (dérivé > composé composé > dérivé)	lexicalisante (phrase > mot)

Par conséquent, l'étymologie populaire intervient non seulement sur des mots simples mais aussi dans le cadre d'un environnement morphologiquement construit. Elle suit les règles des différents procédés morphologiques, comme la lexicalisation, la restructuration, l'attribution du genre et la substitution. Les formations ainsi créées ne contreviennent pas aux règles de la langue et présentent habituellement une grande acceptabilité. Ainsi, elle comporte une dimension et une systématisme morphologiques : elle ne fait certes pas partie des procédures proprement morphologiques, comme la dérivation ou la composition, mais peut avoir une telle fonction. Ce fonctionnement, de nature qui s'apparente aux processus morphologiques, serait compatible avec un ensemble de faits cognitifs avérés, qui se trouvent, pour la plupart, à l'interface entre le niveau conceptuel et le niveau de production et perception du langage.

Étymologie populaire et morpho-sémantique

On peut distinguer trois degrés d'intervention morpho-sémantique sur une échelle ascendante, degrés qui reproduisent l'échelle de sémantisation morphologique, où les allomorphes sont reliés entre eux dans le lexique mental sans différenciation remarquable au niveau du contenu : en dérivation, l'affixe impose son instruction sémantique à la base, et en composition le sens est compositionnel (cf. Corbin 1987) :

- **Renforcement sémantique** : le sens initial de la base n'est pas modifié et dans les dérivés, l'instruction sémantique de l'affixe n'est pas activée, puisqu'il s'agit d'un pseudo-préfixe. L'intervention de l'étymologie populaire contribue à instaurer une transparence morphologique et/ou de renforcement du sens initial. Par conséquent, les bases qui ont fait l'objet d'un tel procédé sont des pseudo-allomorphes de la base initiale, puisqu'ils coexistent avec la base de rapprochement.

-
- (15) (a) GA herpe'ton¹ 'reptile' > GM serpe'to 'reptile'
 (b) TC *perdah* > GM ber'd-aki 'bastonnade'
-

1. Prononciation en GM [erpe'ton].

• Substitution sémantique : même si le référent reste le même, le sens se différencie dans une certaine mesure, vu la substitution sémantique des morphèmes. L'étymologie populaire vise, comme nous l'avons vu, la transparence étymologique, ainsi qu'une meilleure représentativité de la fonction sémantique, grâce à des éléments moins savants, plus productifs et prototypiques (Anastassiadis-Syméonidis 2003b). Ce phénomène rappelle la fonction sémantique de la substitution morphologique, ex. GH *kseko'lo* « décoller, détacher » vs GM *apoko'lo* « décoller, détacher » :

- | | | |
|------|-----|--|
| (16) | (a) | GA phthi'nopo :ron ¹ 'automne' > GM xi'noporo ('fθino 'détériorer, réduire' > 'xino 'verser, partir en vrille') |
| | (b) | GMéd. e'yoklima > GMéd. a'yio klima 'chèvrefeuille' (GA 'aiks 'chèvre' > GMéd. 'ayios 'saint') |

• Resémantisation : le mot qui a fait l'objet d'un procédé relatif à l'étymologie populaire renvoie à un référent différent ou à une notion différente de celle de la base par émergence (a) de l'instruction sémantique des affixes ou (b) de la contribution sémantique des nouvelles bases. Des phénomènes sémantiques de ce type apparaissent en composition aussi bien qu'en dérivation. Les néologismes, où l'étymologie populaire remplit un rôle additionnel par rapport à la transparence qui est celui de la dénomination, entrent dans cette catégorie.

- | | | |
|------|-----|--|
| (17) | (a) | FR <i>gamin</i> > GM xa'mini 'gosse négligé' |
| | (b) | GA 'zeido:ros ² 'fertile' (zei'a 'blé' + 'do:ron 'cadeau') > GM 'zidosros (zo'i 'vie' + 'doro 'cadeau') |

Discussion

Étymologie populaire et phraséologie

Nous avons vu que l'étymologie populaire suit les règles sémantiques de la morphologie, étant donné qu'elle crée des pseudo-allomorphes, substitue des segments, ou encore change le sens de l'unité lexicale initiale, dans un souci de transparence et de dénomination. La plupart des exemples que nous avons présentés ici conduisent à une unité lexicale qui est soit un composé/pseudo-composé, soit un préfixé/pseudo-préfixé, ou encore un suffixé/pseudo-suffixé ou pseudo-allomorphe.

Si, dans certains cas, la forme produite par l'intervention de l'étymologie populaire est analysable sur le plan sémantique de manière relativement directe (*poli'θrona* « fauteuil », *a'drakla* « herbe avec des vertus aphrodisiaques »), dans d'autres un découpage en « morphèmes » plus fantaisiste doit intervenir, afin

1. Prononciation en GM [fθi'noporo].

2. Prononciation en GM ['zidosros].

que l'unité gagne en transparence (*yiuvare'laki* « boulette de viande », *xi'no-poro* « automne »). Pour la première catégorie, si nous considérons l'exemple de *poli'θrona* « fauteuil », qui, rappelons-le, a conduit à des jugements de grécité équivalents à ceux des mots « réellement » grecs (Anastassiadis-Syméonidis & Voga 2010), nous pouvons distinguer un composant sémantique qui repose sur la relation évidente qu'il y a entre l'expérience du locuteur relative à un type de siège dans lequel on est particulièrement confortablement assis, et qui se trouve, en toute logique, assimilé à un siège royal (*'θronos* « trône »). Le préfixe *poli-* vient ensuite renforcer cette impression en revêtant un statut de préfixe d'intensification (Gavriilidou 2015).

Pour l'exemple *sintri'vani* « fontaine », le composant sémantique renvoie de manière assez claire vers la métaphore relative au monde physico-mécanique (*tri'vi* « friction »), renforcé par le trait sémantique « fabriqué par l'homme dans un but décoratif — récréatif ». Cette deuxième catégorie est certes caractérisée par une analyse sémantique moins directe et plus fantaisiste, que l'on retrouve par ex. dans *yiuvare'laki*, où la syllabe *yiuv-* jouerait le rôle du préfixe sans en être un. Cependant, cela ne change rien à la pertinence de l'ensemble conceptuel lié à la parenté perceptive, c'est-à-dire parenté de la forme « arrondie » entre le tonneau et la boulette de viande, sur laquelle se rajoute la réminiscence de la syllabe *yiuv-* rappelant le thème *yir-*, relative à l'idée de « tourner » (ex. *'yiros* « tour », *yi'rizo* « tourner »). Tout ceci en dépit de, et compte non-tenu de la description grammaticale de la langue ou de son histoire (Béguelin 2002, 155).

Dans le « composé » *xi'no-poro*, pour prendre un troisième exemple de cette catégorie, la relation avec le côté décadent de *'xino* (litt. « verser », mais aussi « partir en vrille » dans un usage plus métaphorique) est moins évidente que la relation de préfixation de *poli'θrona* « fauteuil ». Cependant, dans ces trois exemples, nous observons un lien clair avec l'expérience du locuteur, liée non pas uniquement à son expérience de sujet parlant, mais aussi à l'expérience du monde physique qui l'entoure, à l'expérience proprioceptive mais aussi affective et sociale. Le rôle de l'expérience peut être ramené à des notions telles l'iconicité ou l'indexicalité, relatives à la manière dont la réalité psychologique est codée sur le plan cognitif, comme le soutiennent les partisans de la morphologie naturelle (ex. Kilani-Schoch & Dressler 2005). L'opposition entre morphologie prototypique et morphologie non-prototypique (Kilani-Schoch & Dressler 2005) nous semble aussi pertinente afin de décrire le lien récurrent entre l'étymologie populaire et la phraséologie, particulièrement dans leurs aspects métaphoriques.

Lorsque l'on s'intéresse à l'étude de faits morphologiques liés à l'étymologie populaire, nous ne pouvons que repérer les similitudes entre cette dernière et la phraséologie, puisqu'il s'agit dans les deux cas de construire/créer des structures porteuses de sens sur la base de différents domaines de connaissance. À ce propos, le « réseau sémantique », c'est-à-dire les connaissances, images et schémas sur lesquels sont basés les métaphores et les associations qui correspondent à ces

connaissances, images ou schémas (*mappings*, Kövecses 2010), devient ici crucial. Si nous partons du constat de Kövecses (2010, 91) selon lequel la définition d'une métaphore conceptuelle par la formule A est B, dans laquelle un domaine cible est compris en termes d'un domaine source, n'est pas assez précise, il convient de décrire ces associations. Même si elles sont partielles, elles parviennent à constituer un réseau sémantique, en prenant appui sur un aspect précis de l'expérience du monde : le monde naturel, les grandes orientations (HAUT/BAS) et leurs significations culturelles, le corps humain, le monde physique et spirituel, les événements et mythes religieux, etc. On peut ainsi repérer des « métaphores morphologiques », selon l'expression de Nobile (2014, voir aussi Kilani-Schoch & Dressler, 2005). Ces métaphores morphologiques utilisent, parfois de manière agrammaticale, l'affixe, ex. *colissimo* « super-paquet » ou *doctissimo* « super-docteur ». Dans nos exemples le *yiuvare'laki* « boulette de viande et de riz » qui, par le rappel qu'il opère du verbe « tourner » fonctionne comme un intensificateur de l'aspect arrondi du petit tonneau ; ou le suffixe augmentatif - *'akla* de *a'drakla* « herbe aux qualités aphrodisiaques » qui souligne le trait « plus » de surcroît associé à la pseudo-base « homme » (exemples 3c, 12a) ; dans *xa'mini* « gosse négligé » < fr. *gamin* (exemple 12a), le pseudo-suffixe -*'ini* crée l'impression que le mot provient de la famille morphologique de *'xamo* « par terre, en bas ».

Ce qui est important ici n'est pas tellement si ces métaphores morphologiques se construisent de manière qui se prête plus ou moins à une analyse en morphèmes (ou pseudo-morphèmes), mais le fonctionnement à l'intérieur du lexique mental, c'est-à-dire la manière dont se construit le sens. Comme le souligne Kövecses (2010, 91), lorsqu'un domaine-source est appliqué sur une cible, ce ne sont que certains aspects de la cible qui sont mis en exergue (*brought into focus*). Autrement dit, un aspect est mis en relief (*highlighted*) ce qui implique que les autres sont cachés. La sélection de l'aspect qui est mis en relief dépend de la cognition humaine et reflète des similarités non-objectives et non-pré-existantes entre un domaine-source et un domaine-cible. C'est précisément de ce point de vue que ces associations sont pertinentes pour l'étude du « réseau sémantique » présent à l'esprit du sujet (pour le grec, cf. Voga & Anastassiadis-Syméonidis 2020; Voga 2025). Ces similarités constituent la base « expérientielle » des métaphores conceptuelles et sont façonnées par des processus culturels et historiques. Les différents processus morphologiques et morpho-sémantiques décrits tout au long de ce travail remplissent précisément ce but et témoignent de la manière dont ces métaphores font surface dans une langue morphologiquement riche comme le grec.

Conclusions

À partir de toutes ces considérations, nous pouvons formuler les conclusions suivantes concernant les processus liés à l'étymologie populaire en grec :

- Ce sont des processus agissant sur plusieurs niveaux, qui trouvent leur application préférentielle auprès de certaines catégories d'unités lexicales ;
- Que ce soit dans le cadre d'un départ spontané, c'est-à-dire non-conscient, ou d'une intervention consciente, soit verticale, résultant de positions liées à la politique linguistique officielle, soit horizontale, résultant d'un jeu de construction ou reconstruction du savoir lexical au sein de la communauté linguistique, les processus liés à l'étymologie populaire sont profondément ancrés. Ces processus trouvent une bonne partie de leurs sources dans l'expérience du locuteur, telle qu'elle se reflète dans la langue. Ils contribuent grandement à instaurer ou à augmenter la transparence de l'UL qui fait l'objet d'un tel rapprochement ;
- Ces processus n'ont pas uniquement une fonction phonologique et sémantique, mais comportent une dimension morphologique, puisqu'ils interviennent au niveau de la structure, des caractéristiques morphologiques ou des constituants des mots morphologiquement complexes, sur la base des relations paradigmatiques ;
- L'ensemble de ces processus ne peut pas être assimilé au terme étymologie populaire, car nous constatons le phénomène de la parétymologie savante ;
- Ces processus ne peuvent pas être considérés comme des phénomènes aléatoires basés sur l'analogie, mais comme faisant partie d'un mécanisme systématique, étant donné qu'ils sont caractérisés par un certain degré de régularité phonologique et morpho-sémantique. Ils prennent appui sur la forme phonologique de certaines catégories de mots et leur intervention consiste à instaurer un lien entre les unités lexicales concernées sur le plan morpho-sémantique, par analogie au fonctionnement des niveaux morphologique, phonologique et sémantique. Dans ce type de fonctionnement, que nous avons essayé d'esquisser, les notions reliant le niveau conceptuel au niveau linguistique, telles l'iconicité, la prototypicité ou encore celle des métaphores morphologiques devraient être davantage étudiées.

Bibliographie

- AMENTA, Simona et Davide CREPALDI, « Morphological processing as we know it : An analytical review of morphological effects in visual word identification », *Frontiers in Language Sciences*, 3 (2012) : 232 p.
- ANASTASSIADIS-SYMÉONIDIS, Anna, *Néologie d'emprunt. Emprunts directs au français et à l'anglo-américain. Analyse morphophonologique* [Νεολογικός δανεισμός της νεοελληνικής. Άμεσα δάνεια από τη γαλλική και αγγλοαμερικανική. Μορφοφωνολογική ανάλυση], Thessalonique, 2004.

- ANASTASSIADIS-SYMÉONIDIS, Anna, « Que peut-il arriver à une expression figée? », *Cahiers de Lexicologie*, 82.1 (2003a): p. 51-59.
- ANASTASSIADIS-SYMÉONIDIS, Anna, « La structure morphologique du grec moderne et son enseignement / Η μορφολογική δομή της Νέας Ελληνικής και η διδακτική της », *Glossologia*, 15 (2003b): p. 25-34.
- ANASTASSIADIS-SYMÉONIDIS, Anna et Sophia GALANI, « Le suffixe *-ismos* en grec moderne / Το επίθημα -ισμός στη νεοελληνική », *Studies in Greek Linguistics*, 15 (1995): p. 519-529.
- ANASTASSIADIS-SYMÉONIDIS, Anna et Assimakis FLIATOURAS, *La distinction [savant] et [populaire] en grec: définition et classification / Η διάκριση [λόγιο] και [λαϊκό] στην ελληνική γλώσσα: ορισμός και ταξινόμηση*, *Greek linguistics-Proceedings of the 6th International Conference of Greek Linguistics*, Université de Crète, 2005, p. 110-120: <http://www.philology.uoc.gr/conferences/6thICGL/gr.htm> (consulté le 20/10/2025).
- ANASTASSIADIS-SYMÉONIDIS, Anna et Eleni VLETSI, *Faux amis: les amis jurés diachroniques en grec moderne / Ψευδόφιλες μονάδες: οι διαχρονικοί άσπονδοί φίλοι στην ελληνική*, *Colloque panhellénique La didactique de la langue grecque: langue maternelle, seconde et étrangère*, Université de Macédoine Occidentale, Département de l'enseignement Préscolaire, Nymfaio, 2009: <http://linguistics.nured.uowm.gr/Nimfeo2009/praktika/> (consulté le 20/10/2025).
- ANASTASSIADIS-SYMÉONIDIS, Anna et Madeleine VOGA, « Le caractère symbolique des graphèmes du grec moderne / Ο συμβολικός χαρακτήρας γραμμάτων της ΝΕ », *Studies in Greek Linguistics*, 30 (2010): p. 79-97.
- ANTILLA, Raimo, *Analogy*, The Hague/Paris/New York, Mouton, coll. « Trends in linguistics: State of the Art reports », 1977.
- ARONOFF, Marc, *Morphology by itself*, Cambridge, MIT Press, 1994.
- BALDINGER, Kurt, « À propos de l'influence de la langue sur la pensée. Étymologie populaire et changement sémantique parallèle », *Revue de Linguistique Romane*, 37 (1973): p. 241-273.
- BÉGUELIN, Marie-Josée, « Étymologie "populaire", jeux de langage et construction du savoir lexical », *Semen*, 15 (2002): p. 155-172.
- BRUCKER, Charles, *L'étymologie*, Paris, 1988.
- CORBIN, Danielle, *Morphologie dérivationnelle et structuration du lexique. 2 vol.*, Tübingen / Villeneuve d'Ascq, Max Niemeyer Verlag / Presses universitaires de Lille, 1987/1991.
- Dictionnaire du grec standard / Λεξικό της κοινής νεοελληνικής*, Thessaloniki, Institut d'Études Néohelléniques de l'université Aristote de Thessaloniki, 1998.

- FLIATOURAS, Assimakis, Le changement morphologique en grec / Η μορφολογική αλλαγή στην ελληνική γλώσσα, Athènes, Patakis, 2018.
- FÖRSTERMANN, Ernst, « Über deutsche Volksetymologie », *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* [= Kuhns Zeitschrift], 1 (1852) : p. 1-25.
- GAGNÉ, Christina et Thomas SPALDING, « Semantics, concepts, and meta-cognition : Attributing properties and meanings to complex concepts », dans L. BAUER, L. KÖRTVÉLYESSY et P. STEKAUER (éd.), *Semantics of Complex Words*, New York, Springer, 2015, p. 9-25.
- GAGNÉ, Christina et Thomas SPALDING, « Relational diversity affects ease of processing even for opaque English compounds », *The Mental Lexicon*, 9.1 (2014a) : p. 48-66.
- GAGNÉ, Christina et Thomas SPALDING, « Conceptual composition : The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds », dans B.H. ROSS (éd.), *The Psychology of Learning and Motivation*, 19, New York, Elsevier, 2014b, p. 97-130.
- GAVRIILIDOU, Zoé, « The system of intensifying prefixes in Modern Greek », *Linguistica Investigationes*, 37.2 (2014) : p. 240-255.
- GIRAUDO, Hélène et Serena DAL MASO, « The salience of complex words and their parts : Which comes first? », *Frontiers in Psychology* 7 (2016) : p. 17-78.
- GIRAUDO, Hélène et Madeleine VOGA, « Words matter more than morphemes : An investigation of masked priming effects with complex words and nonwords », *Italian Journal of Linguistics studies*, 28.1 (2016) : p. 49-78.
- HATZIDAKIS, Georgios, *Lectures académiques / Ακαδημεικά Αναγνώσματα*, vol. 3, *Linguistique générale*, Athènes, 1915.
- KARANTZOLA, Eleni et Assimakis FLIATOURAS, *Changement linguistique / Γλωσσική αλλαγή*, Athènes, Nisos, 2004.
- KATSOUDA, Georgia et Thanassis NAKAS, *Aspects de la néologie : Mots porte-manteaux et réétymologisation / Όψεις της νεολογίας : Σύμφυση και επανευμολόγηση*, Athènes, Patakis, 2013.
- KILANI-SCHOCH, Marianne et Wolfgang U. DRESSLER, *Morphologie naturelle et flexion du verbe français*, Tübingen, Gunter Narr Verlag Tübingen, 2005.
- KOUTITA-KAIMAKI, Myrto et Assimakis FLIATOURAS, *Blends in Greek dialects : A morphosemantic analysis*, dans A. RALLI, B. JOSEPH et M. JANSE (éd.), *Proceedings of the International Conference on Modern Greek Dialects and Linguistic Theory (MGDLT)*, Patras, University of Patras, 2001, p. 117-130.

- KÖVECSES, Zoltan, *Metaphor. A practical introduction*, New York, Oxford University Press, 2010.
- LONGTIN, Marie-Cécile et Fanny MEUNIER, « Morphological decomposition in early visual word processing », *Journal of Memory and Language*, 53.1 (2005) : p. 26-41.
- MACMAHON, April, *Understanding language change*, New York, Cambridge University Press, 1994.
- MAKRI-TSILIPAKOU, Marianthi, « À propos de quelques formes marquées en grec moderne / Μερικές στιγματισμένες φόρμες της νέας ελληνικής », *Studies in Greek Linguistics*, 6 (1986) : p. 183-196.
- MOÏSIADIS, Théodoros, *Introduction à l'étymologie du grec moyen et moderne / Εισαγωγή στη μεσαιωνική και νεοελληνική ετυμολογία*, Athènes, Ellinika Grammata, 2005.
- MÜLLER, Peter, Ingeborg OHNHEISER, Susan OLSEN et Franz RAINER (éd.), *Word formation : An international handbook of the languages of Europe*, vol. 4, Berlin/ Boston, Mouton de Gruyter, 2016.
- NOBILE, Luca, « Introduction. Formes d'iconicité », *Le français moderne*, 82.1 (2014) : p. 1-45.
- ORR, John, « L'étymologie populaire », *Revue de Linguistique Romane*, 18 (1953) : p. 129-142.
- PALMER, Abram Smythe, *Folk-etymology : A dictionary of verbal corruptions or words perverted in form or meaning, by false derivation or mistaken analogy*, London, George Bell & Sons, 1882.
- PAPANASTASSIOU, Giorgos, *L'Orthographe du grec moderne : Histoire, théorie, application / Νεοελληνική ορθογραφία : Ιστορία, θεωρία, εφαρμογή*, Thessaloniki, Institut d'Études Néohelléniques, 2008.
- PETROUNIAS, Evangelos, *Partie étymologique du Dictionnaire du grec standard*, Thessaloniki, Institut d'Études Néohelléniques, 1998.
- PESETSKY, David, *Russian morphology and lexical theory*, Manuscript, MIT.
- RALLI, Aggeliki, *Morphologie*, Athènes, Patakis, 2005.
- SMITH, Edward et Daniel OSHERSON, « Conceptual combination with prototype concepts », *Cognitive Science*, 8 (1984) : p. 357-361.
- SYMÉONIDIS, Charalambos, *Introduction à l'onomastique grecque / Εισαγωγή στην ελληνική ονοματολογία*, Thessalonique, Kyriakidis, 1992.

- SYMÉONIDIS, Charalambos, *Dictionnaire étymologique des toponymes habités / Ετυμολογικό Λεξικό των Νεοελληνικών Οικωνυμίων* 2 vol., Nicosie/Thessalonique, Centre d'Études du Saint Monastère de Kykkos, 2010.
- RASTLE, Kathy et Matthew DAVIS, « Morphological decomposition based on the analysis of orthography », *Language and Cognitive Processes*, 23 (2008) : p. 942-971.
- RASTLE, Kathy, Matthew DAVIS et Boris NEW, « The broth in my brother's brothel : morpho-orthographic segmentation in visual word recognition », *Psychonomic Bulletin and Review*, 11 (2004) : p. 1090-1098.
- VOGA, Madeleine, Hélène GIRAUDO et Anna ANASTASSIADIS-SYMÉONIDIS, « Differential processing effects within 2nd group Modern Greek verbs », *Lingue e Linguaggio*, 2 (2012) : p. 215-234.
- VOGA, Madeleine et Hélène GIRAUDO, « Word and beyond-word issues in morphological processing », *Word Structure*, 10.2 (2017) : p. 235-254.
- VOGA, Madeleine et Anna ANASTASSIADIS-SYMÉONIDIS, « The role of conceptual metaphors in comprehension of Greek idioms: a self-paced reading study », dans S. MARKANTONATOU et A. CHRISTOFIDOU (éd.), *Multiword expressions in Greek and other languages: from theory to implementation / Πολυλεκτικές εκφράσεις: μελέτες με δεδομένα από την Ελληνική και άλλες γλώσσες*, *Bulletin de terminologie scientifique et de néologismes*, 15, Athènes: Académie d'Athènes, 2020, p. 333-363.
- VOGA, Madeleine, *De l'unité lexicale à l'unité polylexématique. Codage morphologique et métaphorique de la connaissance du monde*, Xanthi/Montpellier, Spanidis et Université de Montpellier Paul-Valéry, 2025.

La traduction en italien du pronom ça sujet

Alberto BRAMATI

Università degli Studi di Milano, Italie

Introduction

Le démonstratif *ça* est un pronom neutre que la tradition grammaticale française a longtemps considéré comme la forme contractée et familière, employée surtout à l'oral, du pronom *cela*¹. En fait, comme les études les plus récentes l'ont montré, *ça* a des propriétés syntactico-sémantiques différentes, qui rendent non seulement son analyse mais aussi sa traduction dans d'autres langues particulièrement complexes. C'est le cas de l'italien, car le pronom *ça* ne peut pas toujours se traduire par l'un des pronoms démonstratifs neutres de cette langue.

Dans cette étude, nous allons nous concentrer sur les propriétés syntactico-sémantiques du pronom *ça* sujet du verbe et sur ses traductions en italien. Ce choix se justifie par la haute fréquence de la fonction sujet dans notre corpus bilingue, composé de 24 textes français contemporains en prose (9 romans, 11 pièces de théâtre et 4 essais en sciences humaines), chacun accompagné de sa traduction italienne² : sur les 2 968 occurrences de *ça* que nous avons collectées, les emplois

1. Voir, entre autres, CHEVALIER *et alii* 2002, 243 ; GREVISSE 1993, 1024 ; RIEGEL, PELLAT, RIOUL 2004, 36 ; WAGNER, PINCHON 1991, 204.

2. Voici les références bibliographiques essentielles des 24 textes constituant notre corpus : ARASSE, D., *On n'y voit rien*, 2013 [2000], tr. it. L. Bianco, *Non si vede niente*, 2013 ; BOILEAU-NARCEJAC, *Celle qui n'était plus (Les diaboliques)*, 2010 [1952], tr. it. F. Di Lella, G. Girimonti Greco, *I diabolici*, 2014 ; CARDINAL, M., *Les mots pour le dire*, 2009 [1975], tr. it. N. Banas, *Le parole per dirlo*, 2000 [1976] ; FABRE, D., *La serveuse était nouvelle*, 2005, tr. it. Y. Melaouah, *La cameriera era nuova*, 2015 ; GAILLY, C., *Un soir au club*, 2006 [2001], tr. it. M. Scotti, *Una sera al club*, 2004 ; GAVALDA, A., *L'échappée belle*, 2012 [2009], tr. it. L. Cisbani, *Il regalo di un giorno*, 2010 ; LAGARCE, J.-L., *Juste la fin du monde*, 1999, tr. it. F. Quadri, « Giusto la fine del mondo », in *Teatro I*, 2009, 43-95 ; LINHARD, R., *L'établi*, 2021 [1978], tr. it. S. Valici, L. Bosio, *Alla catena*, 1978 ; MAUVIGNIER, L., *Loin d'eux*, 1999, tr. it. A. Bramati, *Lontano da loro*, 2009 ; MAUVIGNIER, L., *Ce que j'appelle oubli*, 2011, tr. it. Y. Melaouah, *Storia di un oblio*, 2012 ; MAUVIGNIER, L., *Tout mon amour*, 2012, tr. it. A. Bramati, « Tutto il mio amore », in *Laurent Mauvignier. Théâtre – Teatro*, 2021, 16-173 ; MAUVIGNIER, L., *Une légère blessure*, 2016, tr. it. A. Bramati, « Una ferita leggera », in *Laurent*

de *ça* avec fonction de sujet sont 1 202, soit 45 % du total. Mais si l'on exclut de ce calcul les occurrences du pronom dans des phrases averbales (par ex., *Ça? Oui ça*) et les *ça* éléments d'une expression figée (*ça va?*, *ça y est*, etc.), le pourcentage des *ça* avec fonction de sujet monte à 61 %. Aussi, la fonction sujet est-elle la fonction syntaxique la plus fréquente dans les phrases construites par un prédicat verbal.

Sur la base des données contenues dans notre corpus, nous allons donc présenter dans notre étude, pour chaque emploi de *ça* avec la fonction de sujet, les solutions les plus fréquentes adoptées par les traducteurs professionnels pour rendre cette forme grammaticale en italien. Mais d'abord un aperçu général des propriétés syntactico-sémantiques du pronom démonstratif neutre *ça* et des pronoms démonstratifs correspondants en italien.

Les propriétés syntactico-sémantiques du pronom démonstratif neutre *ça*

À la différence du démonstratif neutre dissyllabique *cela* qui est exclusivement un pronom fort, la forme neutre monosyllabique *ça* a aussi bien les propriétés d'un pronom faible (lorsqu'il a la fonction de sujet¹) que celles d'un pronom fort (dans ses autres fonctions syntaxiques : objet direct, objet prépositionnel, ajout prépositionnel, élément disloqué en tête ou en fin de phrase).

Lorsqu'il est employé comme sujet, le pronom *ça* a donc les mêmes propriétés qu'un pronom personnel clitique :

- *ça* ne peut pas être séparé du verbe par un élément autre qu'un pronom objet clitique ou la particule négative *ne* :

Mauvignier. *Théâtre – Teatro*, 2021, 174-219 ; MICHON, P., *Les Onze*, 2018 [2009], tr. it. G. Girimonti Greco, *Gli Undici*, 2018 ; MODIANO, P., *Dans le café de la jeunesse perdue*, 2014 [2007], tr. it. I. Babboni, *Nel caffè della gioventù perduta*, 2010 ; Mouawad, W., *Incendies*, 2010 [2009], tr. it. C. Gozzi, *Incendi*, 2009 ; NDIAYE, M., *Les serpents*, 2004, tr. it. F. Bevilacqua, C. Montesi, *Le serpi*, 2016 ; POMMERAT, J., *Cendrillon*, 2013, tr. it. C. Gozzi, *Cenerentola*, 2015 ; POMMERAT, J., *La grande et fabuleuse histoire du commerce*, 2012, tr. it. C. Gozzi, « La grande e favolosa storia del commercio », in *La mia cella frigorifera*, *La grande e favolosa storia del commercio*, La riunificazione delle due Coree, 2015, 113-168 ; PONTIUS, J., *À la ligne*, 2022 [2019], tr. it. I. Zagaglia, *Alla linea*, 2022 ; REZA, Y., « Art », in *Théâtre*, 2012 [1994], 187-251, tr. it. F. Di Lella, L. Di Lella, *Arte*, 2018 ; REZA, Y., « L'homme du hasard », in *Théâtre*, 2012 [1995], 5-39, tr. it. C. Spaak, *L'uomo del destino* (trad. inédite ; manuscrit déposé à la SIAE en 2001) ; REZA, Y., *Le dieu du carnage*, 2011 [2007], tr. it. L. Frausin Guarino, E. Marchi, *Il dio del massacro*, 2011 ; SARRAUTE, N., *Pour un oui ou pour un non*, 2003 [1982], tr. it. P. L. Pizzi, *Pour un oui ou pour un non* (trad. inédite pour le spectacle mis en scène en 2021/22 par la Compagnia Orsini) ; SIMENON, G., *L'homme de Londres*, 2019 [1934], tr. it. G. Pinotti, *L'uomo di Londra*, 1999 [1968].

1. « *Ça* peut être considéré comme la forme préverbale de base, avec la variante *ce*, *c'* devant *être* » (PORQUIER 1972, 10). Avec le verbe *être*, le pronom *ça* peut néanmoins s'employer devant les formes composées (*ça a été*), devant les formes simples commençant par une consonne (*ça sera*), devant la particule *ne* ou un autre pronom clitique (*ça n'est pas nouveau* ; *ça m'est égal*), devant les auxiliaires modaux ou aspectuels (*ça peut être*, *ça doit être*, *ça va être*). Voir GREVISSE 1993, 1026.

Ça me plaît beaucoup

Ça ne suffit pas

*Ça, d'ailleurs, devrait être intéressant

- *ça* peut alterner avec les pronoms personnels clitiques sujet et fonctionner comme pronom de reprise d'un sujet nominal ou phrastique, disloqué en tête ou en fin de phrase¹.

Du point de vue sémantique, comme la forme dissyllabique *cela*, le pronom *ça* a d'abord une valeur déictique : il sert à désigner, souvent à l'aide d'un geste, des référents non catégorisés qui se trouvent dans la situation d'énonciation, *i. e.* des référents, animés ou inanimés, qui ne sont pas classés comme des éléments appartenant à une certaine catégorie nominale. Il s'ensuit que, lorsque *ça* a une valeur déictique, la connaissance de la situation d'énonciation est toujours indispensable à la compréhension du message.

Comme la forme *cela*, le pronom *ça* peut aussi avoir une valeur anaphorique, *i. e.* il peut reprendre un élément du contexte de nature et de dimensions variables (un syntagme, une proposition, une phrase, un segment de texte voire le texte tout entier). Cet élément, positionné le plus souvent dans le contexte qui précède, peut se trouver à une distance variable du pronom : d'où, dans certains cas, la difficulté d'établir avec précision l'antécédent de *ça*². Cette difficulté est accentuée par le fait que *ça* ne reprend pas exactement le référent de son antécédent, mais réalise ce que Pierre Le Goffic a appelé une *anaphore médiate*³, *i. e.* une reprise anaphorique qui, en plus du référent de l'antécédent, désigne aussi d'autres informations qu'on peut déduire du contexte. Cette propriété distingue les deux formes *ça* et *cela* : alors que le pronom *cela* reprend toujours un antécédent explicite sans introduire de changements, la forme monosyllabique *ça* enrichit son antécédent d'informations contextuelles et jouit donc d'une autonomie référentielle que la forme dissyllabique ne possède pas :

[...] lors d'un phénomène de reprise, le rapport que les référents de *cela* et *ça* entretiennent avec l'entité introduite au préalable diffère. [...] La différence observée provient de la capacité de *ça* à élargir son référent d'indices contextuels, tandis que *cela* se rapporte strictement à l'énoncé précédent. (Buscail 2010, 138⁴)

Enfin, dans ses emplois déictiques et anaphoriques, le pronom *ça* peut avoir une interprétation indéfinie : employé comme sujet de verbes « évoquant des phénomènes d'ambiance plus ou moins rapportés à des sujets humains » (Riegel, Pellat, Rioul 2004, 451), le pronom *ça* désigne alors un référent « implicite, non spécifié, mais nécessaire » (Le Goffic 1993, 142) qui ne peut être clairement défini :

1. Voir MANENTE 2004, 128-129.

2. PORQUIER 1972, 14.

3. LE GOFFIC 1993, 141-145. Pierre Cadiot (1988) avait désigné la même propriété du pronom *ça* par le terme de « décalage référentiel ».

4. Sur le fonctionnement anaphorique de *ça*, voir aussi GUERIN 2007.

Ça sonne à la porte= quelqu'un
 Ça sent bon= quelque chose

Puisque, tout en étant implicite, le référent du pronom est bien présent dans la réalité, *ça* avec un sens indéfini ne peut pas être substitué par un *il* impersonnel, forme dépourvue de toute valeur référentielle. Là aussi, pour identifier correctement le référent de *ça*, il est indispensable de connaître toutes les informations fournies par le contexte¹.

Les pronoms démonstratifs neutres de l'italien

En italien contemporain, il existe un système bipartite de pronoms démonstratifs : le pronom *questo* désigne un référent proche du locuteur (dans l'espace réel, dans le temps ou dans l'espace du discours) ; au contraire, le pronom *quello* désigne un référent éloigné du locuteur (dans l'espace réel, dans le temps ou dans l'espace du discours²). Ces deux pronoms peuvent être employés, dans la langue orale tout comme dans la langue écrite, aussi bien avec une valeur déictique qu'avec une valeur anaphorique : dans ce cas, l'antécédent du démonstratif peut être un élément nominal, une proposition, une phrase, une portion de texte voire le texte tout entier.

En règle générale, les deux pronoms prennent le genre et le nombre du nom qu'ils substituent. Cependant, au masculin singulier, *questo* et *quello* peuvent avoir un sens neutre.

Outre ces deux formes dissyllabiques, il existe en italien un troisième pronom démonstratif neutre (*ciò*), exclusivement anaphorique, qui appartient au registre soutenu : bien qu'il soit le seul démonstratif monosyllabique, étant surtout employé dans la langue écrite, *ciò* est rarement utilisé pour traduire en italien le pronom *ça*³.

Les trois pronoms démonstratifs de l'italien (*questo*, *quello*, *ciò*) sont tous des pronoms forts.

1. Sur l'emploi de *ça* avec une interprétation indéfinie, voir Manente 2004, 135-136 ; Riegel, Pellat, Rioul 2004, 451. Faute d'exemples dans notre corpus, dans cet article nous ne traiterons pas l'emploi de *ça* « impersonnel », sujet d'un verbe météorologique : pour les propriétés du pronom lorsqu'il est en concurrence avec *il* impersonnel, voir Le Goffic 1993, 142 ; Riegel, Pellat, Rioul 2004, 451 ; Manente 2004, 136-137.

2. En italien, le système des démonstratifs était à l'origine tripartite : à *questo* et *quello* s'ajoutait *codesto*, pronom et adjectif employé pour désigner un référent éloigné du locuteur mais proche de l'interlocuteur. Cette forme n'est aujourd'hui utilisée qu'en Toscane et dans le langage bureaucratique.

3. Les occurrences du pronom *ciò* contenues dans notre corpus sont au total 7, dont 2 seulement comme solution pour traduire un pronom *ça* sujet. Nous y reviendrons dans la conclusion.

La traduction du pronom ça sujet du français vers l'italien

Dans les paragraphes qui suivent, nous allons décrire cinq emplois du pronom démonstratif *ça* avec la fonction de sujet. Notre analyse suivra l'ordre suivant :

1. *ça* déictique
2. *ça* anaphorique intraphrastique (pronom de reprise d'un élément disloqué)
3. *ça* anaphorique interphrastique
4. *ça* indéfini
5. *ça* sujet disloqué

Pour chaque emploi, nous montrerons les solutions les plus fréquentes utilisées dans les textes de notre corpus par les traducteurs professionnels pour exprimer en italien le sens spécifique du pronom français. Commençons par le pronom sujet *ça* avec une valeur déictique.

La traduction du pronom ça déictique

Lorsqu'il est employé avec une valeur déictique, le pronom sujet *ça* se traduit en italien soit par le pronom *questo* (si son référent est proche du locuteur) soit par le pronom *quello* (si son référent est éloigné du locuteur). Les deux pronoms italiens s'accordent en genre et nombre avec le nom italien qui désigne le référent. Étant un pronom exclusivement anaphorique, le démonstratif *ciò* ne peut jamais traduire le pronom *ça* avec une valeur déictique.

Dans l'exemple 1, le pronom *ça* désigne le tableau du peintre Antrios dont causent les deux personnages : comme cet objet est proche des deux interlocuteurs, le sujet *ça* a été traduit en italien par le pronom *questo* (au masculin singulier comme le nom *quadro*, qui traduit le français *tableau*).

I

MARC (<i>désignant l'Antrios</i>). Tu vas le mettre où ?	MARC (<i>indica l'Antrios</i>). Dove lo metterai ?
SERGE. Pas décidé encore. Là. Là ?...	SERGE. Non ho ancora deciso. Qui. O qui ?...
Trop ostentatoire.	Troppo in vista.
MARC. Tu vas l'encadrer ?	MARC. Lo incornicerai ?
SERGE (<i>riant gentiment</i>). Non !... Non, non...	SERGE (<i>ride bonariamente</i>). No !... No, no...
MARC. Pourquoi ?	MARC. Perché ?
SERGE. Ça ne s'encadre pas. (REZA 2012, 216)	SERGE. Questo non si incornicia. (REZA 2018, 45)

La traduction de ça anaphorique pronom de reprise d'un élément disloqué

En tant que forme faible, le pronom *ça* avec fonction de sujet peut être employé comme pronom de reprise d'un élément nominal ou phrastique, disloqué en tête ou en fin de phrase. Lorsque l'élément disloqué est un nom, le pronom *ça*

reste invariable, à la différence des pronoms personnels clitiques de la troisième personne — *il(s)/elle(s)* — qui s'accordent en genre et nombre avec le nom disloqué. Lorsque, en revanche, l'élément disloqué est une proposition, seulement un pronom démonstratif (*ça, ce, cela*) peut être employé comme pronom de reprise du sujet disloqué.

En italien, comme la dislocation du sujet est agrammaticale, le pronom de reprise *ça* disparaît et l'élément disloqué devient le seul sujet du verbe italien¹. Par conséquent, la différence qui existe en français entre la reprise d'un sujet nominal disloqué par un pronom clitique (anaphore directe, interprétation spécifique) et la reprise du même sujet disloqué par le pronom neutre *ça* (anaphore médiate, interprétation générique) ne peut pas s'exprimer linguistiquement en italien. Comme les deux phrases :

Les enfants, ils font du bruit
Les enfants, ça fait du bruit

correspondent en italien à la même phrase,

I bambini fanno rumore

la différence de sens ne pourra s'exprimer que grâce au contexte. Dans le premier cas, le contexte permettra de comprendre que la propriété de faire du bruit caractérise un ensemble défini d'enfants ; dans le deuxième cas, un contexte différent permettra, au contraire, de comprendre que la même propriété caractérise tous les enfants (la classe des enfants en général).

Dans la dislocation à gauche, le sujet déplacé en tête de phrase et repris par le pronom *ça* est presque toujours un élément nominal : à cause de leur longueur, les sujets phrastiques tendent, en effet, à être disloqués à droite. Dans les exemples 2 et 3, le sujet *ça* reprend, respectivement, un nom introduit par un article indéfini avec un sens générique (*une mallette*) et un nom humain pluriel désignant une classe (*les fausses rousse*s²). En italien, puisque la dislocation du sujet est agrammaticale, le pronom *ça* disparaît et l'élément nominal disloqué à gauche devient le seul sujet du verbe. Par conséquent, la virgule qui souvent sépare en français l'élément disloqué du pronom de reprise est effacée dans le texte italien.

2

ANNETTE. Un homme doit être libre de ses mains... Je trouve. Même une mallette, ça me gêne. (REZA 2011, 68)	ANNETTE. Un uomo deve avere le mani libere. Secondo me. Anche una ventiquattr'ore mi dà fastidio. (REZA 2011, 79)
--	--

1. L'agrammaticalité de la dislocation du sujet en italien s'explique par le fait que les pronoms personnels sujet (*io, tu, lui*, etc.) ne sont pas de clitiques. Le dispositif de la dislocation est, en revanche, possible en italien pour toutes les fonctions syntaxiques qui admettent une reprise par un pronom personnel clitique (objet direct, objets prépositionnels introduits par *a, di, da*, objets et ajouts à valeur locative).

2. En tant que pronom de reprise d'un élément nominal, *ça* peut aussi reprendre un nom abstrait dérivé d'un verbe (*la mort*), un nom de masse (*la graisse*), un mot quelconque dans une référence méta-linguistique (*demos*).

3

Finalement, les Anglais ont raison. Ils sont pragmatiques. Ils n'hésitent pas. Un seul mot: *hair*. Comme ça, pas de problème, pas de questions inutiles: le cheveu et le poil, pour une Anglaise, c'est pareil. D'ailleurs, il paraît qu'elles sont toutes rousses. Sûrement des vraies rousses. En Angleterre, *les fausses rousses*, à mon avis, **ça** n'existe pas. (ARASSE 2013, 98-99)

Alla fine hanno ragione gli inglesi, che sono pragmatici. Non esitano, loro: una sola parola, *hair*. Così, niente problemi, niente questioni inutili: per un'inglese, capello e pelo sono la stessa cosa. Del resto, pare, loro sono tutte rosse. Sicuramente tutte vere rosse. Secondo me in Inghilterra **le finte rosse** non esistono. (ARASSE 2013, 76)

La modalité de traduction est la même lorsque, plus rarement, l'élément disloqué à gauche est une proposition. Dans l'exemple 4, l'élément disloqué est représenté par un verbe à l'infinitif (*vendre*), repris par le pronom sujet *ça*. En italien, le pronom de reprise disparaît et l'infinitif correspondant (*vendere*) devient le seul sujet du verbe (*improvvisarsi*).

4

RENÉ. Si tu vexes les gens, t'es cuit. C'est des études américaines sur les comportements, c'est imparable, Franck, on te l'a dit.
MICHEL. Tu crois encore que **vendre ça** s'improvise! Vendre c'est une suite de petits détails... Si tu les respectes tu vends, sinon tu vends pas... (POMMERAT 2012, 24-25)

RENÉ. Se irrito le persone sei finito... Sono ricerche americane sui comportamenti, non si sbagliano Franck, te l'abbiamo detto.
MICHEL. Pensi ancora che **vendere** si improvvisi! Vendere è un susseguirsi di piccoli dettagli... Se li rispetti vendi, se no non vendi... (POMMERAT 2016, 132)

Les mêmes noms qui sont susceptibles d'être disloqués à gauche peuvent aussi être disloqués à droite. Dans l'exemple 5, le pronom sujet *ça* reprend le nom abstrait *mariage*, dérivé du verbe *se marier*. En italien, comme pour les dislocations à gauche, le pronom de reprise disparaît et le nom abstrait correspondant (*matrimonio*) devient le seul sujet du verbe (*consumare*).

5

SERGE. Tu ne trouves pas qu'Yvan a beaucoup maigri?
MARC. Elle aussi.
SERGE. **Ça** les ronge **ce mariage**. (REZA 2012, 216)

SERGE. Non trovi che Yvan sia molto scui-pato?
MARC. Anche lei.
SERGE. Li sta consumando, **questo matrimonio**. (REZA 2018, 44)

Ce qui caractérise les dislocations à droite, c'est surtout le détachement d'éléments phrastiques. Dans notre corpus, le pronom sujet *ça* reprend cinq types d'éléments phrastiques:

- les complétives sujet [*que* Ph]
- les infinitives sujet [(*de*) Vinf]
- les relatives substantives sujet [*ce* + (*qui* + *que* + *dont* + prép *quoi*) + Ph]
- des propositions introduites par la conjonction *quand* [*quand* Ph]
- des propositions introduites par la conjonction *si* [*si* Ph].

Nous allons en donner deux exemples.

Dans l'exemple 6, le pronom *ça* reprend la complétive sujet disloquée à droite (*que tu appelles*). Comme dans les dislocations à gauche, le pronom de reprise disparaît en italien et la complétive correspondante devient le seul sujet du verbe de la phrase (ici la locution verbale *fare piacere*¹).

6

Le téléphone sonne, René va répondre. [...]	Il telefono suona. René risponde. [...]
RENÉ. Allô oui... Ça me fait plaisir que tu appelles... (POMMERAT 2012, 8)	RENÉ. Pronto sì... Mi fa piacere che mi chiami ... (POMMERAT 2016, 117)

Dans l'exemple 7, le pronom *ça* apparaît trois fois comme sujet de la locution verbale *vouloir dire*: alors que dans ses deux premières occurrences, *ça* reprend une infinitive disloquée à droite (*aimer trop* ; *être loin de toi*), dans sa troisième occurrence *ça* reprend une proposition introduite par la conjonction *quand* (*quand tu n'es pas là*). Dans ce cas, bien que la même construction soit en principe possible en italien (*non so cosa vuol dire quando tu non ci sei più*), le traducteur a préféré normaliser la syntaxe, en substituant la la conjonction *quando* par la conjonction *che* (*non so cosa vuol dire che non ci sei più*). Dans les trois cas, le pronom de reprise *ça* disparaît en italien.

7

VOIX DE WAHAB. Je voulais te dire que cette nuit, mon cœur est plein d'amour, il va exploser. Par-tout on me dit que je t'aime trop ; moi, je ne sais pas ce que ça veut dire aimer trop , je ne sais ce que ça veut dire être loin de toi , je ne sais pas ce que ça veut dire quand tu n'es plus là . (MOUAWAD 2010, 38-39)	VOIX DE WAHAB. Volevo dirti che stanotte il mio cuore colmo d'amore sta esplodendo. Tutti mi dicono che ti amo troppo ; io non so cosa vuol dire amare troppo , non so cosa vuol dire essere lontano da te , non so cosa vuol dire che non ci sei più . (MOUAWAD 2009, 35)
---	---

La traduction du pronom *ça* anaphorique interphrastique

Lorsque l'antécédent du pronom sujet *ça* se trouve dans le contexte qui précède (le cas le plus fréquent) ou dans le contexte qui suit (plus rarement), on peut distinguer trois différentes formes d'anaphore :

- 1. *ça* a pour antécédent un élément nominal (nom, pronom, syntagme nominal) ;
- 2. *ça* a pour antécédent un élément phrastique (proposition ou phrase) ;
- 3. *ça* a pour antécédent une portion de texte.

À chaque forme d'anaphore correspondent des solutions différentes pour traduire le pronom *ça* en italien. Voyons en détail les trois cas.

1. La locution verbale *faire plaisir* appartient à une classe sémantiquement homogène de verbes qui, dans le langage familier, expriment la réaction psychologique provoquée par le sujet phrastique sur la ou les personne(s) désignées par l'objet du verbe (ici le clitique accusatif *me*). Cette classe de verbes a souvent pour sujet un élément phrastique disloqué à droite, repris par le pronom *ça* (*Ça contrarie Luc que Ph; Ça m'ennuie de Vinf*). Voir LE GOFFIC 1993, 249-250.

Le pronom sujet *ça* a pour antécédent un élément nominal

Lorsque le pronom *ça* a pour antécédent un élément nominal positionné dans le contexte qui précède, sa traduction en italien par un pronom démonstratif (*questo/quello*) est impossible. Les solutions que les traducteurs adoptent sont alors généralement trois : 1) l'omission du pronom (lorsque le contexte permet d'identifier sans ambiguïté le sujet nominal non exprimé du verbe italien) ; 2) la répétition de l'élément nominal qui a la fonction d'antécédent ; 3) les structures [*è una cosa che*] ou [*è qualcosa che*], où le sujet sous-entendu du verbe *essere* est l'élément nominal antécédent du pronom *ça* tandis que son attribut est un nom ou un pronom neutre (*cosa/qualcosa*) suivi d'une phrase relative. Les deux exemples qui suivent permettent de voir ces trois possibilités de traduction.

Dans l'exemple 8, le pronom *ça*, qui apparaît deux fois comme sujet du verbe *finir* (*ça finit, ça finit*), a pour antécédent le syntagme nominal *l'amour*. En italien, le premier pronom a été omis : le sujet sous-entendu du verbe italien *finire* est, en effet, clairement le syntagme nominal *l'amore*. Le deuxième pronom, en revanche, a été traduit par le même syntagme nominal qui a la fonction d'antécédent (*l'amore*) : la présence, dans la complétive qui précède, d'un syntagme nominal qui pourrait, du moins au premier abord, être interprété comme le sujet du verbe (*la fine*) rend en effet problématique l'omission du vrai sujet. Dans les deux cas, la traduction de *ça* par le démonstratif *questo* serait impossible.

8

Ah oui, un rêve d'enfant, c'est ça pour moi l'amour, la nostalgie de quand j'étais petite et que rêvais d'amour, un rêve de vieillard qui s'invente sa vie à défaut de l'avoir vécue. Tu sais ce que c'est toi? [...] Et je peux te dire, moi, comment ça finit, parce que je sais que la fin — écoute! ça finit alors qu'on est encore complètement dans l'euphorie [...]

(MAUVIGNIER 2016, 28)

Ah sì, un sogno infantile, è questo per me l'amore, la nostalgia di quando ero piccola e sognavo l'amore, il sogno di un vecchio che si inventa la vita per non averla vissuta. Lo sai tu che cos'è? [...] E posso dirtelo io, come finisce, perché io so che la fine — ascolta! l'amore finisce quando si è ancora completamente immersi nell'euforia [...]

(MAUVIGNIER 2021, 199-201)

Dans l'exemple 9, le pronom *ça* a pour antécédent le syntagme nominal *le silence*, qui apparaît trois fois dans le contexte précédent. Dans ce cas, à cause de la distance qui sépare le pronom de son antécédent, le sujet du verbe italien ne peut pas être omis. Probablement pour éviter une quatrième répétition du même syntagme, le traducteur n'a pas opté ici pour la lexicalisation du pronom (*Il silenzio ti cade addosso*), mais plutôt pour sa traduction par la structure [*è qualcosa che*], où le sujet sous-entendu du verbe *essere* est l'antécédent nominal de *ça* (*il silenzio*). Dans cet exemple aussi, la traduction du pronom français par le démonstratif *questo* serait impossible.

À se taire aussi on savait bien ce qui se passait, dans nos têtes à tous les deux, à Jean surtout que je regardais, lui dont le regard se perdait sur ses chaussures, avec ses pieds qui grattaient le sol [...] et les tapotements secs qui ne couvriraient rien du silence dans lequel on était, car maintenant je sais qu'on était dans le silence, que le silence ce n'est pas quand il n'y a pas de bruit, ou qu'on n'entend rien, ou quand tout seul on se repose. Ça tombe dessus. (MAUVIGNIER 1999, 33)	Anche tacendo lo sapevamo bene quello che succedeva nelle nostre teste, soprattutto in quella di Jean, io lo guardavo, Jean, mentre il suo sguardo si perdeva sulle scarpe, con i piedi che grattavano il terreno [...] e i colpetti secchi che non coprivano niente del nostro silenzio, perché adesso lo so che stavamo nel silenzio, che il silenzio non è quando non c'è rumore, o quando non si sente niente, o quando da soli ci si riposa. È qualcosa che ti cade addosso. (MAUVIGNIER 2008, 25)
---	---

Le pronom sujet ça a pour antécédent un élément phrastique

Lorsque le pronom sujet *ça* a pour antécédent un élément phrastique (proposition ou phrase) positionné dans le contexte qui précède, les solutions pour sa traduction en italien changent. Outre que par son omission, quand le contexte permet d'identifier sans ambiguïté le segment de texte qui a la fonction d'antécédent, le pronom *ça* peut se traduire par un pronom démonstratif neutre (*questo/quello*), par un syntagme contenant le nom neutre *cosa* ou par une lexicalisation. Les exemples qui suivent permettent de voir ces différentes solutions.

Dans l'exemple 10, les deux occurrences du pronom *ça* reprennent la même phrase *lui, le croit*, d'abord sous la forme d'une complétive (*Que lui le croie ne te gêne pas?*), ensuite sous la forme d'une infinitive (*Non. Si de le croire lui fait plaisir*). Dans les deux cas, comme le sujet des formes verbales correspondantes (le verbe *infastidire* et la locution verbale *fare piacere*) peut être identifié sans aucune ambiguïté, le traducteur a opté pour l'omission du pronom. Ici, à la différence des exemples 8 et 9, le pronom *ça* aurait pu aussi être traduit par le démonstratif neutre *questo* (*Questo non ti infastidisce? / Se questo gli fa piacere*).

10

MARC. Désormais, notre ami Serge fait partie du Gotha des grands amateurs d'art. YVAN. Mais non!... MARC. Bien sûr que non. À ce prix-là, on ne fait pas partie de rien, Yvan. Mais, lui, le croit. YVAN. Ah oui... MARC. Ça ne te gêne pas? YVAN. Non. Si ça lui fait plaisir. (REZA 2012, 203)	MARC. Ormai il nostro amico Serge fa parte del Gotha dei grandi appassionati d'arte. YVAN. Ma no!... MARC. Certo che no. A quel prezzo non fai parte di niente, Yvan. Ma lui ne è convinto. YVAN. Ah sì... MARC. Non ti infastidisce? YVAN. No. Se gli fa piacere (REZA 2018, 25)
--	---

C'est bien le démonstratif neutre *questo* qui a été utilisé pour traduire en italien le pronom *ça* dans l'exemple 11. Cet exemple montre que le choix de traduire *ça* par un pronom neutre peut dépendre non seulement de la nécessité d'éviter toute ambiguïté dans l'identification de l'antécédent de *ça*, mais aussi de raisons mélodico-rythmiques. Ici, l'antécédent de *ça* étant la phrase qui précède (*nous avons notre ignorance de l'autre en commun*), le traducteur aurait pu omettre sans inconvé-

nient le sujet du verbe italien (*noi due abbiamo in comune l'ignoranza dell'altro e può bastarci, no?*): s'il a opté pour la traduction de *ça* par le pronom *questo*, c'est seulement, dans ce cas, pour améliorer le rythme de la phrase.

II

Peu m'importe ce qu'il y a derrière tes sourires et tes gestes, peu m'importe ce que tu comprends de moi et ce que je comprends de toi, nous avons notre ignorance de l'autre en commun et **ça** peut nous suffire, non ?

(MAUVIGNIER 2016, 24)

Non mi importa cosa c'è dietro i tuoi sorrisi e i tuoi gesti, non mi importa cosa capisci di me e cosa capisco io di te, noi due abbiamo in comune l'ignoranza dell'altro e **questo** può bastarci, no ?

(MAUVIGNIER 2021, 195)

Une troisième solution pour traduire en italien le sujet *ça* lorsqu'il a pour antécédent un élément phrastique positionné dans le contexte qui précède est un syntagme nominal contenant le nom neutre *cosa*: *la cosa, questa cosa, è una cosa che*.

Dans l'exemple 12, le pronom *ça* a pour antécédent la phrase qui précède (*on ne m'avait rien trouvé*): dans ce cas, au lieu de traduire *ça* par un pronom neutre (*Questo mi aveva riempito di gioia*), solution tout à fait possible, le traducteur a opté pour le syntagme nominal *la cosa*.

12

J'avais pris une demi-journée pour un spécialiste de la prostate et la visite gratuite de la sécurité sociale, et on ne m'avait rien trouvé. **Ça** m'avait rempli de joie pendant deux jours, le temps d'une méchante gueule de bois.

(FABRE 2005, 34)

Mi ero preso una mezza giornata per andare da uno specialista della prostata a fare la visita gratuita che passa la mutua, e non mi avevano trovato niente. **La cosa** mi aveva riempito di gioia per due giorni, il tempo di farmi passare una sbronza coi fiocchi.

(FABRE 2015, 21)

Dans l'exemple 13, les deux occurrences du pronom *ça* ont pour antécédent la même phrase (*Les hommes sont tellement accrochés à leurs accessoires*). Dans ce cas, au lieu de traduire *ça* par un pronom neutre (*Questo li sminuisce*) ou par un syntagme contenant le nom neutre *cosa* (*Questa cosa li sminuisce*), le traducteur a opté pour la structure [*è una cosa che*], où le sujet du verbe *essere* est la phrase qui représente en français l'antécédent de *ça*.

13

ANNETTE. Les hommes sont tellement accrochés à leurs accessoires... **Ça** les diminue. **Ça** leur enlève toute autorité. Un homme doit être libre de ses mains... Je trouve.

(REZA 2011, 68)

ANNETTE. Gli uomini sono talmente attaccati ai loro accessori... **È una cosa che** li sminuisce. **Che** toglie loro ogni autorità... Un uomo deve avere le mani libere. Secondo me.

(REZA 2011, 79)

Dans l'exemple 14, enfin, le pronom *ça* a pour antécédent une relative substantivée (*ce qu'il disait*). Dans ce cas, comme le pronom neutre *questo* aurait produit une séquence cacophonique avec la traduction du groupe [*ce que*] (*quello che questo significava*), le traducteur a opté pour une lexicalisation: le pronom *ça* a donc été traduit par un syntagme nominal qui résume le contenu de la relative substantivée (*quello che Philippe mi diceva* → *le sue parole*).

14

Combien de temps il m'a fallu pour comprendre qu'au contraire c'était le bon moment pour que j'entende ce qu'il disait, pour entendre ce que **ça** voulait dire contre moi, ce qu'il aurait aimé que j'entende [...]
(MAUVIGNIER 2000, 72-73)

Quanto tempo ci ho messo a rendermi conto che invece era il momento giusto per capire quello che Philippe mi diceva, per capire quello che **le sue parole** significavano contro di me, quello che lui avrebbe voluto che io capissi [...]
(MAUVIGNIER 2008, 50)

Le pronom sujet ça a pour antécédent une portion de texte

Lorsqu'enfin le pronom sujet *ça* a pour antécédent une portion de texte, sa traduction par le pronom neutre *questo*, qui désigne ce qui précède immédiatement le pronom, n'est possible que si la cette portion peut être interprétée comme une unité conceptuelle. C'est pourquoi, dans la plupart des cas, les traducteurs choisissent d'autres solutions : le syntagme nominal pluriel *queste cose*, le syntagme pronominal *tutto questo* ou, tout simplement, le pronom indéfini *tutto*. Les quatre exemples qui suivent illustrent ces différentes solutions de traduction.

Dans l'exemple 15, le pronom sujet *ça* a pour antécédent trois phrases appartenant au contexte précédent (*t'es neuf, t'y connais rien, t'as une bonne mentalité*) : dans ce cas, étant donné que les trois phrases peuvent être interprétées comme une unité conceptuelle, le pronom *ça* a été traduit en italien par le pronom neutre *questo*.

15

MICHEL. Donc, voilà ce qui a été essentiel pour nous, dans notre décision de te demander de venir travailler avec nous : t'es neuf, t'y connais rien... Et d'après moi t'as une bonne mentalité... C'est **ça** qui a convaincu les autres quand je leur ai parlé de toi...
(POMMERAT 2012, 6)

MICHEL. Ecco cosa è stato fondamentale, per decidere di chiederti di venire a lavorare con noi: sei nuovo, non conosci niente. E secondo me hai una buona mentalità... È **questo** che ha convinto gli altri quando gli ho parlato di te.
(POMMERAT 2015, 115-116)

Dans l'exemple 16, le pronom *ça* permet au moins deux interprétations : si l'on considère qu'il désigne exclusivement la dernière phrase du discours du personnage nommé Homme 1 (*tu étais à l'autre bout du monde*), *ça* peut être traduit en italien par le pronom neutre *questo* (*questo mi fa male*) ; si, au contraire, on considère, comme l'a fait le traducteur, que *ça* reprend le discours entier du personnage, la traduction par le pronom *questo* s'avère impossible car ce pronom ne désignerait que la dernière phrase, celle qui précède le pronom. Pour désigner le discours entier du personnage, le traducteur a donc opté pour une solution différente, à savoir le syntagme nominal pluriel *queste cose*.

16

HOMME 1: Si, dis-moi... je te connais trop bien: il y a quelque chose de changé... Tu étais toujours à une certaine distance. de tout le monde, du reste. mais maintenant avec moi... encore l'autre jour, au téléphone. tu étais à l'autre bout du monde. **ça** me fait de la peine, tu sais... (SARRAUTE 2003, 24)

FRANCO. No, ti conosco troppo bene: qualcosa è cambiato. Tu, è vero, ti sei sempre tenuto ad una certa distanza da tutti... anche con me. ancora l'altro giorno, al telefono, sembravi fuori dal mondo... lontano... e **queste cose** mi fanno male lo sai...
(SARRAUTE 2021, 3-4)

Dans l'exemple 17, le personnage nommé Yvan raconte à ses amis Serge et Marc le démêlé qu'il a eu avec sa fiancée Catherine après un coup de fil de sa mère à propos de l'organisation de leur mariage. Dans ce cas, le pronom sujet *ça* désigne non seulement le fait spécifique d'« être sur le carton » mais, plus en général, tout ce qui concerne la participation de la belle-mère de Catherine à la fête de mariage : *ça* reprend donc non seulement le texte qui précède mais aussi d'autres informations qu'on peut déduire du contexte. Ici, le pronom *ça* a été traduit en italien par le syntagme pronominal *tutto questo*, qui a un sens plus large que le syntagme nominal *queste cose* utilisé dans l'exemple 16.

17

YVAN. [...] au revoir, elle raccroche, Catherine, qui était à côté de moi, mais qui ne l'avait pas entendue, me dit, qu'est-ce qu'elle dit, je lui dis, elle ne veut pas être sur le carton avec Yvonne et c'est normal, je ne parle pas de ça, qu'est-ce qu'elle dit sur le mariage, rien, tu mens, mais non Cathy je te jure, elle ne veut pas être sur le carton avec Yvonne, rappelle-la et dis-lui que quand on marie son fils, on met son amour-propre de côté, tu pourrais dire la même chose à ta belle-mère, ça n'a rien à voir, s'écrit Catherine, c'est moi, moi, qui tiens absolument à sa présence, pas elle, la pauvre, la délicatesse même, si elle savait les problèmes que *ça* engendre, elle me supplierait de ne pas être sur le carton, rappelle ta mère [...] (REZA 2012, 220)

YVAN. [...] arrivederci, riaggancia, Catherine, che era accanto a me, ma che non l'aveva sentita, mi dice che ha detto, le dico non vuole stare sulle partecipazioni insieme a Yvonne ed è normale, non parlavo di questo, che ha detto del matrimonio, niente, non dire bugie, ma no, Cathy, te lo giuro, non vuole stare sulle partecipazioni insieme a Yvonne, richiamala e dille che una madre, quando il figlio si sposa, mette da parte l'orgoglio, be', potresti dire la stessa cosa alla moglie di tuo padre, non c'entra niente, urla Catherine, sono io, io, che tengo assolutamente alla sua presenza, non lei, poverina, che è la discrezione in persona, se sapesse dei problemi che **tutto questo** sta creando, mi supplicherebbe di non essere sulle partecipazioni, richiama tua madre [...] (REZA 2018, 52)

Dans l'exemple 18, enfin, le pronom *ça* désigne tout le discours où le locuteur (*lui*) parle des bruits qu'il entend dans sa tête et des choses qui lui tordent le ventre, en décrivant les symptômes et les pensées qui l'accompagnent dans ces moments difficiles. Dans ce cas, pour désigner le contenu entier de cette séquence, le pronom *ça* a été traduit par le pronom indéfini *tutto*.

18

Et pourtant, l'avant-dernière fois c'est lui qui a parlé, qui m'a dit, tu comprends papa, au travail, ces bruits, des fois des choses sans nom qui vous tordent le ventre et qu'on ne sait pas nommer, alors qui nous mangent encore plus à cause de ça, qu'on ne sait pas les nommer et pourquoi c'est sur nous seuls qu'on croit qu'elles s'abattent, ces espèces de bruits. Et la tête qui résonne et frappe encore des heures après que *ça* s'est calmé comme c'est venu, par hasard on dirait. (MAUVIGNIER 1999, 80)

Eppure, la penultima volta è stato lui a parlare, a dirmi, capisci papà, al lavoro, a volte, quei rumori, quelle cose senza nome che ti torcono lo stomaco e non sai come chiamare e ti mangiano vivo proprio per questo, perché non sai come chiamarle e perché credi che si abbattano solo su di te, tutte quelle specie di rumori. E la testa che risuona e che batte ancora per ore dopo che **tutto**, così com'è venuto, si è calmato, per caso sembrerebbe. (MAUVIGNIER 2008, 60)

La traduction du pronom *ça* indéfini

Avec une valeur déictique ou anaphorique, le sujet *ça* peut prendre, dans certains contextes, un sens indéfini : dans ces cas, le référent du pronom, tout en étant toujours présent (d'où l'impossibilité de substituer *ça* avec l'impersonnel *il*), reste implicite. Pour traduire en italien le pronom *ça* avec un sens indéfini, il existe principalement trois solutions : le pronom indéfini *tutto*, le syntagme nominal avec un sens indéfini *le cose* et une lexicalisation. Voyons un exemple pour chacune de ces trois solutions.

Dans l'exemple 19, le pronom *ça* désigne de façon vague la situation personnelle de la femme à qui s'adresse le locuteur : dans ce cas, le traducteur a choisi de rendre le sujet *ça* par le pronom indéfini *tutto*.

19

J'ai fait ma tête de vrai benêt, tout ignorant, et puis je lui ai dit en regardant la belle femme que **ça** finirait sans doute par s'arranger, comme les fois d'avant. (FABRE 2005, 43)

Ho fatto la faccia da beota, di quello che non sa niente, e poi guardando la bella donna le ho detto che di sicuro alla fine si aggiustava **tutto**, come le altre volte. (FABRE 2015, 27)

Le syntagme nominal *le cose* est, en revanche, la solution qui a été choisie dans l'exemple 20 : dans ce cas, le pronom *ça*, sujet du verbe *se passer*, désigne tout ce qui peut arriver dans la vie. Le même sens indéfini aurait pu être exprimé, comme dans l'exemple précédent, par le pronom indéfini *tutto* (*voglio che tutto vada bene*).

20

MAURICE. C'est quoi ta vraie motivation à toi pour être là ?

FRANCK. Je sais pas... Je me suis installé avec quelqu'un y a pas longtemps. Avant j'avais que moi à m'occuper, maintenant on est deux... Et j'ai envie que **ça** se passe bien, j'ai envie de gagner ma vie. (POMMERAT 2012, 7)

MAURICE. Qual è la vera motivazione per cui sei qui ?

FRANCK. Non lo so... Sono andato a vivere da poco con qualcuno. Prima dovevo pensare solo a me stesso, adesso siamo in due. E voglio che **le cose** vadano bene, voglio guadagnarmi da vivere. (POMMERAT 2015, 117)

La troisième solution est la lexicalisation, à savoir la traduction du pronom *ça* par un syntagme nominal. S'agissant d'exprimer un sens indéfini, le choix du nom dépend non seulement du contexte mais aussi, dans une large mesure, de l'interprétation du traducteur.

Dans l'exemple 21, le pronom *ça* désigne la brève période de temps où les parents peuvent promener leur enfant dans une poussette : le pronom *ça* a été donc traduit en italien par un nom de temps (*quel momento*).

21

ANNETTE. Mon mari n'a jamais été un père à poussette!...

VÉRONIQUE. C'est dommage. C'est merveilleux de promener un enfant. **Ça** passe si vite. (REZA 2011, 22)

ANNETTE. Mio marito non è mai stato di quelli che portano a spasso i figli nel passeggino!...

VÉRONIQUE. Peccato. È bellissimo portare a spasso un bambino. **Quel momento** dura così poco. (REZA 2011, 24)

La traduction du pronom *ça* sujet disloqué en tête ou en fin de phrase

Le pronom *ça* avec la fonction de sujet peut être disloqué en tête ou en fin de la phrase : dans ce cas, *ça* a toutes les propriétés d'un pronom fort tandis que le pronom de reprise — un autre pronom *ça* ou, devant certaines formes du verbe *être*, sa variante *ce* — garde les propriétés d'un pronom faible. Du point de vue de sa traduction, puisqu'en italien — nous l'avons déjà vu — la dislocation du sujet est agrammaticale¹, le pronom de reprise (*ça* ou *ce*) disparaît tandis que le pronom *ça* disloqué devient le seul sujet du verbe italien.

Comme le pronom sujet faible, le sujet *ça* disloqué en tête ou en fin de la phrase peut avoir aussi bien une valeur déictique qu'une valeur anaphorique. Le pronom *ça* déictique se traduit normalement en italien, selon la position de son référent par rapport au locuteur, par les pronoms démonstratifs *questo* ou *quello*, accordés en genre et nombre avec le nom qui désigne le référent. C'est ce qu'on voit dans l'exemple 22, où *ça* apparaît trois fois comme sujet disloqué avec une valeur déictique, toujours repris par sa variante *ce*, sujet du verbe *être* : alors que, dans la première occurrence, *ça* est disloqué à droite (*c'est quoi, ça?*), dans les deux autres *ça* est disloqué à gauche (*ça, c'est une robe*). En désignant un *manichino* (nom masculin singulier) proche de l'interlocuteur, le premier *ça* a été traduit par le pronom italien *quello* ; en revanche, en désignant un *vestito* (nom masculin singulier) proche du locuteur, le deuxième *ça* a été traduit par le pronom italien *questo*. Plus rarement, si le contexte permet de reconstruire sans ambiguïté le sujet du verbe italien, *ça* avec une valeur déictique peut être omis (c'est le cas de la troisième occurrence de *ça* disloqué dans l'exemple 22).

22

LA BELLE-MÈRE (*désignant le mannequin*) Je te demande c'est quoi **ça**?

LE PÈRE (*jouant l'incompréhension*). Ça?

LA BELLE-MÈRE. (*insistant encore*). Oui ça!

LE PÈRE. Ah ça.

LA BELLE-MÈRE (*exaspérée*). Tu me prends vraiment pour une idiote toi!

LE PÈRE. Ah mais **ça**, c'est une robe c'est tout!

LA BELLE-MÈRE. **Ça** c'est une robe c'est tout?!
(POMMERAT 2013, 28)

LA MATRIGNA (*indicando il manichino*) Ti ho chiesto che cos'è **quello**?

IL PADRE (*facendo finta di non capire*). Quello???

LA MATRIGNA. (*insistendo ancora*). Sì, quello!

IL PADRE. Ah, questo.

LA MATRIGNA (*esasperata*). Mi prendi proprio per un'idiota!

IL PADRE. Ah ma **questo** è solo un vestito!

LA MATRIGNA. È solo un vestito?!
(POMMERAT 2015, 21)

Les mêmes solutions peuvent être utilisées lorsque le sujet *ça* disloqué a une valeur anaphorique. Dans l'exemple 23, le pronom *ça*, disloqué à gauche et repris par sa variante *ce*, reprend la phrase prononcée par Maurice (*vendre un produit comme le nôtre en ce moment, c'est pas évident*) : puisque cet antécédent est positionné juste avant le pronom, *ça* a été traduit en italien par le démonstratif neutre *questo*.

1. Voir § 3.2 ci-dessus.

23

MAURICE. En plus, vendre un produit comme le nôtre en ce moment, c'est pas évident.

ANDRÉ. **Ça** c'est sûr... Vaudrait mieux leur vendre des trucs à bouffer, aux gens. (POMMERAT 2012, 37)

MAURICE. Oltretutto vendere un prodotto come il nostro in questo momento, non è così facile.

ANDRÉ. **Questo** è certo... A questa gente sarebbe meglio vendergli della roba da mangiare. (POMMERAT 2016, 144)

Comme le *ça* déictique, le *ça* anaphorique disloqué en tête ou en fin de phrase peut ne pas être traduit lorsque le contexte rend superflue l'expression du sujet du verbe italien. C'est le cas de l'exemple 24, où les quatre occurrences du pronom *ça* disloqué à gauche (*Ça c'est vrai!*) ont pour antécédent les phrases prononcées par Annette: dans ce cas, au lieu de traduire *ça* par le pronom neutre *questo*, le traducteur a opté pour l'omission du sujet du verbe *essere*, obtenant ainsi un texte plus rapide et léger (*Questo è vero!* → *È vero!*).

24

MICHEL. Je n'ai absolument pas tué ce hamster!

ANNETTE. C'est pire. Vous l'avez laissé tremblant d'angoisse dans un milieu hostile. Ce pauvre hamster a dû être mangé par un chien ou par un rat.

VÉRONIQUE. **Ça** c'est vrai! **Ça** c'est vrai!

MICHEL. Comment ça, **ça** c'est vrai!

VÉRONIQUE. **Ça** c'est vrai. Qu'est-ce que tu veux! C'est affreux ce qui a dû arriver à cette bête. (REZA 2011, 47)

MICHEL. Non ho assolutamente ucciso il criceto!

ANNETTE. Peggio. L'ha abbandonato tremante di paura in un ambiente ostile. Quel povero criceto sarà stato divorato da un cane o da un topo.

VÉRONIQUE. È vero! È vero!

MICHEL. Come sarebbe, è vero?

VÉRONIQUE. È vero! Ha ragione! A quella povera bestia sarà successo qualcosa di orribile. (REZA 2011, 54)

Les mêmes solutions de traduction peuvent être employées lorsque le sujet *ça* est disloqué à droite. C'est ce qu'on voit dans l'exemple 25, où le pronom sujet *ça* disloqué à droite a pour antécédent un syntagme nominal désignant un être humain (*un demi-campagnard un peu ahuri*): dans ce cas, le pronom *ça* exprime le mépris de l'Embauteur¹ pour les candidats ouvriers qu'il sélectionne. La même connotation péjorative a été exprimée en italien par l'emploi du démonstratif masculin singulier *questo*.

25

Quant à moi, j'avais l'air suffisamment accablé pour faire un candidat ouvrier insoupçonnable. Monsieur l'Embauteur a dû penser: «Tiens, un demi-campagnard un peu ahuri, c'est bon, **ça**: ça ne fera pas d'histoires.» Et il m'a donné mon bon pour la visite médicale. (LINHART 2021, 16)

Quanto a me, avevo l'aria sufficientemente affranta per costituire un candidato operaio insospettabile. Il Signor Collocatore avrà pensato: «Guarda, un mezzo contadino un po' tonto, buono **questo**, non farà storie.» E mi ha dato il mio tagliando per la visita medica. (LINHART 1979, 12)

1. Lorsqu'il désigne un être humain, le pronom *ça* exprime «quelque mouvement affectif, qui peut aller du mépris à la tendresse» (GREVISSE 1993, 1024). La même remarque dans CHEVALIER *et al.* 2002, 243.

Conclusion

Les données traductologiques que nous venons de présenter sont schématiquement résumées dans le tableau récapitulatif qui suit :

Tableau 1

I	LA TRADUCTION DU PRONOM ÇA SUJET					
1	ça déictique	<i>questo</i>	<i>quello</i>			
2	ça anaphorique intraphrastique — pronom de reprise d'un élément disloqué					
2.1	<i>X, ça V</i>	Ø				
2.2	<i>ça V..., X</i>	Ø				
3	ça anaphorique interphrastique					
3.1	<i>ça = SN</i>	Ø	<i>lexicalisation</i>	<i>è una cosa che è qualcosa che</i>		
3.2	<i>ça = proposition ça = phrase</i>	Ø	<i>questo neutre quello neutre</i>	<i>questa cosa la cosa</i>	<i>è una cosa che</i>	<i>lexicalisation</i>
3.3	<i>ça = portion de texte</i>	<i>questo neutre</i>	<i>queste cose</i>	<i>tutto questo</i>	<i>tutto</i>	
4	ça indéfini	<i>tutto</i>	<i>le cose</i>	<i>lexicalisation</i>		
5	ça sujet disloqué					
5.1	<i>ça, c'est / ça, ça V</i>	<i>questo</i>	<i>quello</i>	Ø		
5.2	<i>c'est.ça / ça V...ça</i>	<i>questo</i>	<i>quello</i>	Ø		

Sans jamais oublier que « tout corpus ne représente que lui-même », comme le disait Maurice Gross, les données que nous avons collectées nous permettent quelques remarques conclusives :

1. *ça* pronom de reprise d'un élément nominal ou phrastique disloqué en tête ou en fin de phrase (anaphore intraphrastique) ne se traduit pas en italien : la dislocation du sujet étant en italien agrammaticale, le pronom disparaît toujours dans les traductions ;
2. la traduction de *ça* par un pronom démonstratif neutre (*questo/quello*) est impossible en italien lorsque *ça* reprend un élément nominal extraphrastique et lorsqu'il a un sens indéfini : dans ces deux cas, le traducteur doit recourir à d'autres solutions (l'omission du pronom, une lexicalisation ou les structures [è (*una cosa* + *qualcosa*) *che*] pour l'anaphore interphrastique ; le pronom indéfini *tutto*, le syntagme nominal *le cose* ou une lexicalisation pour *ça* avec un sens indéfini) ;
3. les démonstratifs neutres *questo* et *quello* traduisent *ça* surtout lorsque le pronom français reprend une proposition, une phrase ou une plus large portion

de texte (anaphore interphrastique) ; les mêmes pronoms peuvent aussi traduire le pronom *ça* disloqué en tête ou en fin de phrase lorsque son référent est neutre ;

4. Le pronom *ça* déictique se traduit par les pronoms démonstratifs *questo* ou *quello*, accordés en genre et nombre avec le nom italien qui désigne le référent de *ça*.

Deux remarques pour conclure : 1) bien qu'il soit le seul démonstratif italien monosyllabique, le pronom neutre *ciò* apparaît rarement dans notre corpus : le pronom *ça* étant surtout employé dans la langue orale familière, un pronom tel que *ciò*, utilisé surtout à l'écrit, peut difficilement rendre en italien la forme française ; 2) la variété de solutions que les traducteurs italophones ont adoptées pour traduire les différents emplois de *ça* sujet montre bien la nécessité d'une définition précise du sens du pronom dans un certain contexte avant de proposer une traduction quelconque. L'étude d'autres corpus bilingues permettra sans aucun doute d'enrichir la liste des traductions possibles en italien pour chaque emploi du pronom sujet *ça* en français, enrichissant ainsi la liste des solutions que nous avons proposées ici.

Bibliographie

- BLANCHE-BENVENISTE, Claire *et al.*, *Pronom et syntaxe. L'approche pronominale et son application au français*, Paris, Selafl, 1987.
- BLANCHE-BENVENISTE, Claire *et al.*, *Le français parlé. Études grammaticales*, Paris, Éditions du C.N.R.S., 1990.
- BLASCO-DULBECCO, Mylène, *Les dislocations en français contemporain. Étude syntaxique*, Paris, Champion, 1999.
- BRAMATI, Alberto, « Le problème de la traduction en italien du pronom *ça* valenciel dans *Loin d'eux* de Laurent Mauvignier », *Synergies Italie*, 6 (2010) : p. 83-93.
- BRAMATI, Alberto, *Grammatica contrastiva per i traduttori dal francese*, Milano, Edizioni Libreria Cortina, 2025.
- BUSCAIL, Laurie, « Qu'est-ce que *ça* veut dire ? », *Linx*, 62-63 (2010) : p. 135-151.
- CADIOT, Pierre, « De quoi *ça* parle ? À propos de la référence de *ça*, pronom-sujet », *Le Français Moderne*, année 56, 3/4 (oct. 1988) : p. 174-192.
- CALABRESE, Andrea, « I dimostrativi: pronomi e aggettivi », dans L. RENZI, G. SALVI, A. CARDINALETTI (éd.), *Grande grammatica italiana di consultazione*, Bologna, il Mulino, vol. 1, 2001 [1988], p. 631-645.

- CARLIER, Anne, « “Les gosses ça se lève tôt le matin” : l’interprétation générique du syntagme nominal disloqué au moyen de *ce* ou *ça* », *Journal of French Language Studies*, 6(02) (1996) : p. 133-162.
- CHEVALIER, Jean-Claude *et al.*, *Grammaire du français contemporain*, Paris, Larousse, 2002 [1964].
- DA MILANO, Federica, « Dimostrativi, aggettivi e pronomi », *Enciclopedia dell’italiano*, Roma, Istituto dell’Enciclopedia Italiana, 2010, p. 373-374.
- GREVISSE, Maurice (1993), *Le bon usage*, 13^e éd. refondue par A. GOOSSE, Paris, Louvain-La-Neuve, Duculot.
- GUERIN, Emmanuelle, « “Moi je comprends pas et ça ça m’intrigue”. Ce que peut nous apprendre la double occurrence de *ça* dans les énoncés oraux présentant un phénomène de dislocation », *Linx*, 57 (2007) : p. 27-37.
- LE GOFFIC, Pierre, *Grammaire de la phrase française*, Paris, Hachette, 1993.
- LEFEUVRE, Florence, « Les anaphores résomptives en *c’*, *cela*, *ça* et *ceci* dans *Juste la fin du monde* et *Derniers remords avant l’oubli* de Jean-Luc Lagarce », dans C. DOQUET et É. RICHARD (éd.), *Les représentations de l’oral chez Lagarce*, Paris, L’Harmattan, 2012, p. 111-133.
- MAILLARD, Michel, « “Un zizi, ça sert à faire pipi debout!”. Les références génériques de *ça* en grammaire de phrase », dans G. KLEIBER (éd.), *Rencontres avec la généricité*, Paris, Klincksieck, 1987, p. 157-206.
- MANENTE, Mara, « Une analyse syntaxique du pronom *ça* », *Annali di Ca’ Foscari* XLIII, 1-2 (2004) : p. 125-149.
- PORQUIER, Rémy, « L’emploi de *ça* en français parlé », *Le français dans le monde*, 91 (1972) : p. 9-16.
- RIEGEL, Martin, PELLAT, Jean-Christophe, RIOUL, René, *Grammaire méthodique du français*, Paris, PUF, 2004.
- SERIANNI, Luca, *Italiano*, con la collaborazione di A. CASTELVECCHI, Milano, Garzanti, 1997 [1988].
- WAGNER, Robert Léon, PINCHON, Jacqueline, *Grammaire du français classique et moderne*, Paris, Hachette, 1991.

Enseigner la politesse chinoise aux apprenants francophones : représentations et difficultés d'appropriation

Guo JING
INaLCO/ PLIDAM

Introduction

Dans l'enseignement des langues, l'accent est souvent mis sur les compétences linguistiques, bien que l'on accorde également une certaine attention à la culture des peuples qui les parlent. L'apprenant perçoit la langue étrangère comme un nouveau code linguistique, auquel il associe des connaissances parfois stéréotypées sur la culture du pays où cette langue est utilisée. Cependant, en raison d'explications souvent superficielles ou de l'absence totale de certains éléments du code culturel en classe, il développe une compétence partielle, insuffisante pour interagir pleinement avec les locuteurs natifs. Il ne parvient alors pas à comprendre certains comportements des natifs dans leurs échanges quotidiens. La politesse, composante essentielle de ce code culturel, a un impact direct sur la qualité des interactions, et son degré de maîtrise s'avère déterminant.

Au département d'études chinoises de l'Inalco, comme dans la plupart des centres de formation, l'enseignement met l'accent sur la langue standard, dans un registre généralement neutre, voire formel. Les formules de politesse de base, telles que « Bonjour! » « Au revoir! » « Merci! » et « De rien! » sont introduites dès les premières étapes de l'apprentissage. Cependant, en Chine, ces expressions ne sont utilisées que dans des contextes restreints, souvent très formels, et ne conviennent pas aux échanges entre personnes proches. À la fin de la première année de licence, les apprenants parviennent à se débrouiller avec les formules de base leur permettant de communiquer avec des locuteurs chinois en France, même s'ils les emploient fréquemment de manière inappropriée. Or, à partir de la

deuxième année de licence, nombre d'entre eux partent étudier un semestre ou une année en Chine. Ils se retrouvent alors confrontés à des situations où ils doivent interagir avec des personnes devenues très proches, dans des contextes beaucoup plus familiers. Par ailleurs, les différences culturelles entre la Chine et la France influencent les actes de politesse, qui ne se manifestent pas toujours de la même manière dans les deux pays. Certains comportements chinois peuvent sembler impolis aux Français, tandis que certains usages français sont perçus comme tels en Chine. Nos apprenants sont-ils conscients de ces différences culturelles? Quelle est leur compréhension de la politesse chinoise? Quelles difficultés rencontrent-ils dans l'apprentissage de ces formes de politesse? Et comment pouvons-nous les aider? Dans cet article, nous commencerons par analyser les représentations que nos apprenants ont des formules de politesse chinoises, avant de présenter une méthode d'enseignement que nous avons conçue pour favoriser leur appropriation de ces expressions.

La politesse chinoise

La politesse linguistique désigne « tous les aspects du discours qui sont régis par des règles, et dont la fonction est de préserver les caractères harmonieux de la relation interpersonnelle » (Kerbrat-Orecchioni 1996, 51). Selon la théorie de la « face » de Goffman (1982), chaque individu se présente comme ayant adopté une certaine ligne de conduite, qu'il choisit en fonction de ce qu'il perçoit comme les attentes sociales propres à la situation dans laquelle il se trouve, et qu'il se doit de maintenir sous peine de « perdre la face ». La préservation de la face, ainsi que les règles de politesse qui en découlent, relèvent d'un principe universel. Certains actes peuvent être considérés comme menaçants pour la face (*Face Threatening Acts*, FTA, Brown & Levinson 1987), mais certains peuvent aussi être valorisants (Kerbrat-Orecchioni 1992).

Les conceptions de la face et de la politesse varient d'une culture à l'autre. Les locuteurs occidentaux et orientaux partagent de nombreux principes de politesse, tels que la générosité, la modestie et la bienveillance, et adoptent des comportements similaires dans des situations correspondantes. Cependant, malgré ces ressemblances pragmatiques, les différences culturelles entre la Chine et la France se manifestent à tous les niveaux en matière de politesse. Il est souvent dit que les Chinois, notamment les touristes chinois en Europe, sont réputés pour leurs « mauvaises manières ». Ils sont fréquemment critiqués pour manger, boire et parler bruyamment, sans se préoccuper de leur environnement. Ces comportements heurtent les codes de conduite occidentaux. Si certains de ces comportements relèvent de l'éducation, d'autres sont davantage liés à la culture, en particulier aux usages (Traverso 2000), c'est-à-dire la fréquence d'un comportement dans une situation donnée. Un comportement peut être poli sans pour autant

être conforme aux usages. Par exemple, dans un magasin en Chine, un client français pourrait dire « Bonjour! » ou « Au revoir! » à un vendeur chinois, ce qui serait poli mais inhabituel. Inversement, un comportement peut être conforme aux usages mais non poli, ce qui peut entraîner des malentendus. Par exemple, face à un problème sérieux, comme une erreur commise, les Chinois ont tendance à sourire pour désamorcer la situation, un comportement souvent mal perçu par les Français. De même, certains usages français peuvent être considérés comme impolis en Chine, tels que l'appel d'un interlocuteur plus âgé par son prénom, se moucher ou s'embrasser en public, ou encore offrir un cadeau en présence d'autres personnes.

En Chine, la politesse est perçue comme un critère de moralité, un moyen d'exprimer intention, conduite et éthique sociales (Bi Jiwan 毕继万 1996). Elle se traduit en chinois par le terme 礼貌 *lǐmào*, dont le caractère 礼 *lǐ* revient au cœur de la pensée confucéenne. Ce dernier renvoie aux rites, aux normes sociales ainsi qu'aux comportements conformes aux convenances. L'objectif des 礼 *lǐ* est de créer une société où chaque individu se comporte en fonction de sa position sociale (Gu Yueguo 顾曰国 1992). Sous cette influence, les Chinois privilégient le collectivisme et minimisent l'importance de l'individu. La valeur personnelle et l'espace privé ne sont guère valorisés, car ils sont souvent perçus comme égoïstes. Ainsi, les Chinois, souvent réservés et discrets, n'expriment que rarement leurs sentiments en public. Pour préserver leur propre face ainsi que celle de leur interlocuteur, ils évitent les conflits et s'efforcent de maintenir une ambiance harmonieuse.

Le concept de famille, qui englobe l'idée d'appartenance d'un individu à une entité collective dans tous les aspects de sa vie (foyer, quartier, école, entreprise, ville, région, État), occupe une place centrale dans la pensée chinoise. Dans cette microsociété, chacun cherche à s'identifier comme un membre contribuant au développement collectif. C'est pourquoi les Chinois manifestent facilement le désir de se rapprocher des personnes faisant partie de leur structure sociale (amis, voisins, collègues, connaissances), voire même d'inconnus. Dans les conversations, des sujets très personnels tels que le salaire, l'âge ou la vie familiale sont fréquemment abordés, montrant ainsi leur souci pour l'autre. Par exemple, au lieu de dire simplement « Bonjour! » un Chinois peut demander à son interlocuteur s'il a déjà mangé à l'heure du repas, ou s'il va sortir s'il le croise quittant l'immeuble, ou encore « Où allez-vous? / Que faites-vous? » Ces questions, qui peuvent sembler intrusives ou impolies aux yeux d'un Occidental, sont en réalité des salutations et un signe de rapprochement.

Quant à l'utilisation des termes de politesse en chinois, elle présente certaines caractéristiques spécifiques que nous aborderons ci-après.

— Utilisation asymétrique des termes de politesse

Comme mentionné précédemment, les Chinois ont tendance à s'auto-dévaloriser pour manifester modestie et respect envers leurs interlocuteurs (Gu Yueguo 顾曰国 1992). En général, plus la différence de statut hiérarchique et la distance relationnelle entre les locuteurs est grande, plus les formules de politesse sont employées. On observe ainsi une utilisation asymétrique de ces termes, où le locuteur appartenant à une classe sociale inférieure s'adresse à celui de la classe supérieure avec davantage de marques de respect. Par exemple, les jeunes utilisent plus de termes de politesse envers les aînés, les employés montrent plus de respect envers leurs employeurs et les directeurs, etc.

— Utilisation des termes d'adresse

En fonction des statuts sociaux et des relations hiérarchiques des locuteurs, les Chinois utilisent différents termes d'adresse pour définir, établir ou ajuster les relations avec leurs interlocuteurs, qu'il s'agisse de les renforcer ou de créer une distance (Gu Yueguo 顾曰国 1992). Cette tendance se manifeste souvent par l'utilisation fréquente de termes de parenté pour s'adresser aux autres, même à des inconnus, tels que « oncle », « tante », « sœur » ou « frère ». Sous l'influence de la culture chinoise, des cultures occidentales et des idéologies politiques propres à chaque époque, les termes d'adresse en chinois présentent une grande complexité. Certains d'entre eux évoluent dans leur signification avec le temps, ce qui peut provoquer des ambiguïtés et des malentendus sur ce qui est (im)poli (Kádár 2011), comme c'est le cas du terme 小姐 *xiǎojiě* (mademoiselle) (Zhang 2022), que nous allons aborder plus loin.

— Utilisation des adoucisseurs syntaxiques et lexicaux, ainsi que des moyens paraverbaux tels que l'intonation, l'accentuation, le rythme, les sourires et les gestes

Il est fréquent que les Chinois n'utilisent pas explicitement de formules ou de marqueurs de politesse tels que « s'il vous plaît », « vous », « merci », « désolé », etc. En revanche, ils transmettent souvent des messages de manière polie et compréhensible par tous en employant des mots ou des structures syntaxiques qui expriment la sympathie, la sincérité et la gratitude, ou simplement en adoptant un ton moins directif, accompagné d'un sourire. Il est maintenant bien établi que plus les relations entre les locuteurs sont proches, moins ils utilisent de formules de politesse. Ainsi, au sein d'une famille, il est très courant que les membres n'utilisent pas de termes de salutation, de remerciement, de demande ou d'excuse (Meng 2001). La politesse est alors transmise de manière indirecte et implicite.

Ces quelques explications montrent l'importance de bien connaître les usages de la politesse dans un contexte donné. En se fondant uniquement sur sa propre conception occidentale de la politesse, un Français peut se forger une image très négative des Chinois face à certains de leurs comportements.

Les représentations de la politesse chinoise chez les apprenants

Qu'entendent-ils par politesse chinoise? Quelles sont les difficultés qu'ils rencontrent dans l'apprentissage de la politesse chinoise? Afin de répondre à ces questions, nous avons distribué, lors de la première semaine de formation en septembre 2023, un questionnaire aux étudiants de deuxième année de licence de l'Inalco. Dans ce questionnaire, cinq points sur la politesse chinoise ont été présentés: 1) les titres d'adresse; 2) les actes de salutation; 3) les formules de remerciement; 4) les actes de réception de cadeaux; 5) des usages conventionnels «impolis». Pour chacun de ces points, nous proposons des situations avec des exemples concrets et demandons aux apprenants de choisir les formules appropriées ou inappropriées.

Voici les principaux résultats:

Le transfert des normes de la politesse française

Les réponses (total = 31) recueillies illustrent diverses facettes du phénomène du système interculturel de comportement des apprenants en langues étrangères, tel que décrit par Wang Jianqin 王建勤 (1995, *the learner's system of intercultural behavior*). Ce système comportemental culturel fait référence au processus psychologique par lequel les apprenants, influencés par leur culture d'origine et la culture cible, développent leur propre système culturel. C'est dans ce cadre que se manifeste le phénomène de transfert culturel, lequel consiste en l'application, souvent inconsciente, des concepts issus de la culture maternelle pour comprendre et adapter les comportements communicatifs dans la culture cible. Ce transfert se traduit fréquemment par l'adoption de modes de pensée, de codes de conduite et de fonctions pragmatiques propres à la langue maternelle.

Prenons l'exemple des formules de remerciement. En chinois, les remerciements peuvent être classés en deux catégories: directs et indirects. Les remerciements directs sont similaires à ceux du français, tels que «Merci!» ou «Merci beaucoup!» En revanche, les remerciements indirects utilisent des expressions comme «Vous avez travaillé dur!» ou «Je suis désolé pour les problèmes que j'ai causés!» pour transmettre la gratitude de manière plus implicite. Les Chinois privilégient souvent ces formes indirectes de remerciement. Comme mentionné précédemment, il est courant en Chine que «Merci» ne soit pas utilisé dans les interactions avec des connaissances. À l'inverse, les Français sont plus directs et l'utilisent fréquemment (Meng 2001). Dans notre étude, 61% (n = 31) des apprenants ont choisi «Merci!» pour répondre à un compliment, tandis que 70,9% ont opté pour «Merci beaucoup» pour remercier un ami ayant aidé au déménagement. Bien que ce taux diminue dans les contextes plus intimes, l'utilisation de remerciements directs reste notable: 25,8% des apprenants ont choisi «Merci!» pour décrire la réaction d'une mère dont le fils a réparé un téléphone, et 35,4%

pour celle d'une mère recevant un cadeau de sa fille. Certains ont même choisi 太好了, 我得给你多少钱? *Tài hǎo le, wǒ děi gěi nǐ duōshǎo qián?* (C'est super, combien je te dois?), pourtant inacceptable dans ce contexte en Chine, à hauteur de 19,3 %.

Le transfert culturel est également perceptible chez les apprenants lorsqu'ils identifient des comportements jugés impolis. Certains usages considérés comme impolis en France sont conventionnels en Chine, et *vice versa*. Nous avons demandé aux apprenants de sélectionner les comportements qu'ils jugent impolis en Chine. Nous avons observé une grande confusion entre usages conventionnels et impolis. Par exemple, 64,5 % des apprenants estiment qu'il est impoli de ne pas dire « Bonjour! » ni « Au revoir! » dans un magasin, alors que ce comportement est en réalité courant en Chine. En revanche, des actes tels que se ronger les ongles et se moucher en public, jugés impolis en Chine, n'ont été identifiés comme tels que par une minorité d'apprenants (29 % pour se ronger les ongles, 35,4 % pour se moucher en public). Comme nous l'avons mentionné précédemment, en Chine, les termes d'excuse sont rarement utilisés pour exprimer une erreur ou un malentendu ; les Chinois préfèrent des moyens indirects comme les sourires, la modestie ou les gestes sympathiques. Ainsi, lorsqu'une personne marche sur le pied de quelqu'un et sourit sans s'excuser, cela n'est pas considéré comme impoli si le sourire est accompagné de gestes amicaux. En revanche, selon les normes françaises, ce comportement est perçu comme agaçant, et la majorité des apprenants (80,6 %) le jugent impoli.

D'autres exemples de transfert culturel ont été observés dans l'utilisation des termes d'adresse et des formules de salutation : 小姐 *xiǎojiě* (mademoiselle, 70 %), 夫人 *fūren* / 女士 *nǚshì* (madame, 35,4 %), 你好 *nǐ hǎo* (Bonjour, 77,42 %), 先生 *xiānsheng* (monsieur, 35,4 %), et 请问 *qǐngwèn* (S'il vous plaît, je voudrais savoir..., 70,9 %). Par souci de concision, nous ne détaillons pas ces observations dans cet article.

Le processus du filtrage des normes de la politesse chinoise

Un autre phénomène observé dans le comportement interculturel des apprenants est le processus de filtrage culturel (Zhang Zhanyi 张占一 1990 ; Wang Jianqin 王建勤 1995). Selon Zhang Zhanyi 张占一 (1990), lorsqu'un apprenant communique, il s'exprime en adoptant les formes propres à la langue cible et à sa culture. Cependant, lorsqu'il reçoit des informations dans cette langue, il les « digère » à travers le prisme de sa langue et de sa culture maternelles. En raison des grandes différences entre les contextes culturels et historiques chinois et français, il est souvent difficile pour l'apprenant de trouver des correspondances entre les pratiques des deux peuples. L'apprenant interprète alors la culture chinoise à travers des concepts issus de la culture française, ce qui peut « déformer » la culture chinoise, la rendant parfois méconnaissable et créant ainsi une barrière à la compréhension (Zhang Zhanyi 张占一 1990).

Prenons l'exemple d'une formule de remerciement. Pour remercier un ami qui a aidé lors d'un déménagement, nous avons proposé aux apprenants une phrase parmi les réponses possibles : 今天你累坏了吧! *Jīntiān nǐ lèi huài le ba!* (Aujourd'hui tu es épuisé, n'est-ce pas!) Cette expression, tout aussi courante que « Merci beaucoup! » en chinois, est fréquemment utilisée pour remercier quelqu'un qui a fourni un effort physique. Elle est souvent employée entre personnes proches. Cependant, pour les apprenants français, des termes comme « fatigué » ou « épuisé » évoquent des connotations très négatives. Ainsi, au lieu de percevoir cette expression comme une marque de reconnaissance ou un message bienveillant, beaucoup la jugent inacceptable, voire insultante ; seulement 19,3 % l'ont choisie comme une formule de remerciement, comme en témoigne un étudiant :

Si on nous dit qu'on est épuisé ou très fatigué après l'avoir aidé, cela signifie qu'on est faible, ou incompetent, c'est très insultant.

En effet, selon les apprenants, certaines formules semblent tellement exagérées et peu naturelles qu'ils n'osent pas les employer. Prenons l'exemple des termes d'adresse 小姐 *xiǎojiě* (mademoiselle) et 美女 *měinǚ* (belle femme). Au début du xx^e siècle, le mot 小姐 *xiǎojiě* (mademoiselle) était un titre honorifique utilisé pour désigner une jeune femme riche, célibataire et de bonne éducation. Après 1949, ce terme a été perçu comme un produit de la « conscience de classe exploiteuse » et a acquis une forte connotation péjorative. Suite à la réforme et à l'ouverture du pays, il a commencé à être réintroduit pour désigner les jeunes femmes en général. Cependant, dans les années 1990, ce terme, qui avait été populaire dans les années 1980, a encore changé de sens, évoquant désormais une connotation liée aux « prestataires de services sexuels ». Cela a créé un fossé dans la représentation des jeunes femmes dans la société chinoise contemporaine (Shao Jingmin 邵敬敏 2009) et les Chinois évitent aujourd'hui de l'utiliser.

美女 *měinǚ* (belle femme) est un autre terme moderne utilisé pour s'adresser aux jeunes femmes. À l'origine, il désignait une femme d'une grande beauté. Ce terme est apparu en Chine continentale à la fin des années 1990 et au début des années 2000 pour désigner les femmes en général (Shao Jingmin 邵敬敏 2009). Il combine les éléments « belle » et « femme ». Aujourd'hui, le sens de « belle » s'est élargi, et le degré de beauté qu'il suggère est devenu flou, voire effacé. Dans la plupart des contextes, toutes les jeunes femmes peuvent être qualifiées de 美女 *měinǚ* (belle femme) (Ji Xinxin 姬欣欣 2020).

Dans notre étude, nous avons demandé aux apprenants de choisir le terme approprié pour demander l'heure à une jeune passante. Il en ressort que, pour les apprenants, 美女 *měinǚ* (belle femme) paraît trop exagéré dans cette situation : seulement 22,5 % d'entre eux l'ont choisi, tandis que la grande majorité (70,9 %) préférèrent 小姐 *xiǎojiě* (mademoiselle), qu'ils trouvent plus transparent et naturel en français.

Le même phénomène se retrouve dans l'utilisation des formules pour réagir à un compliment. Nous avons demandé aux apprenants de choisir une réponse appropriée lorsqu'un ami peintre leur a fait un compliment sur un tableau qu'ils ont peint. Seulement 9,6 % ont choisi la formule « Ce n'est pas beau, c'est moche ! » En Chine, ce genre de réponse est pourtant parfaitement acceptable pour exprimer la modestie, une qualité hautement valorisée (Gu Yueguo 顾曰国 1992).

L'adaptation culturelle aux normes de la politesse chinoise

Outre les phénomènes de transfert et de filtrage culturel, les résultats de notre étude permettent également d'observer l'état d'adaptation culturelle des apprenants aux normes de la politesse chinoise. L'adaptation culturelle (Wang Jianqin 王建勤 1995) désigne le processus par lequel l'apprenant s'ajuste continuellement à la culture de la langue cible tout au long de son apprentissage. Ce processus comporte plusieurs étapes. La première, celle de la non-adaptation, correspond à une phase où l'apprenant ne parvient pas encore à s'intégrer dans la culture cible et doit s'appuyer sur les stratégies et les connaissances de sa langue maternelle. Ensuite vient la phase de début d'adaptation, où il acquiert des notions de communication et des connaissances culturelles élémentaires. La troisième étape, dite d'adaptation de base, se caractérise par une meilleure maîtrise de la langue et une réduction des erreurs linguistiques. Enfin, la dernière étape est celle de l'adaptation complète, où l'apprenant atteint presque la même compétence communicative que les locuteurs natifs (Wang Jianqin 王建勤 1995).

Une question de notre questionnaire illustre ce phénomène. Nous avons demandé aux apprenants : « Un enseignant demande le nom d'un étudiant, comment peut-il lui adresser la parole ? » Les options proposées et les réponses obtenues sont les suivantes :

请问，您贵姓？ *Qǐngwèn, nín guì xìng?* (S'il vous plaît, quel est votre nom ? 22,58 %)
 先生，您叫什么？ *Xiānsheng, nín jiào shénme?* (Monsieur, quel est votre nom ? 35,48 %)
 请问，你叫什么？ *Qǐngwèn, nǐ jiào shénme?* (S'il te plaît, comment tu t'appelles ? 70,97 %)
 你叫什么？ *Nǐ jiào shénme?* (Comment tu t'appelles ? 41,94 %)

Les résultats révèlent que les apprenants reconnaissent que les locuteurs de statuts sociaux différents s'adressent les uns aux autres de manière distincte. Dans notre exemple, l'enseignant, ayant un statut supérieur à celui de l'étudiant, n'a pas besoin d'utiliser le terme honorifique « vous ». Beaucoup d'apprenants semblent avoir saisi ce code, puisqu'ils ont majoritairement choisi des formules employant le terme d'adresse « tu ». Cependant, en raison des influences de la culture française, les apprenants tendent à associer une demande d'information avec une expression de politesse telle que « s'il te plaît ». Ainsi, la grande majorité des apprenants ont

combiné le terme d'adresse « tu » avec l'expression « s'il te plaît » (« S'il te plaît, comment tu t'appelles? ») représentant 70,97 %.

D'autres exemples confirment cette tendance. Nous observons que les apprenants commencent à distinguer les formules de politesse entre contextes familiaux et publics. C'est le cas pour l'utilisation des termes de salutation. Bien qu'ils aient tendance à conserver le terme « Bonjour! » dans toutes les situations en raison du transfert culturel, lorsqu'ils saluent un ami, ils optent sans hésitation pour des formules plus informelles, telles que 你干什么去啊? *Nǐ gàn shénme qù a?* (Que vas-tu faire?), choisies par 90,32 % des apprenants.

Par ailleurs, certains apprenants ayant déjà séjourné en Chine parviennent à percevoir une certaine subtilité dans l'utilisation des formules de politesse par les locuteurs natifs chinois. Voici quelques témoignages illustrant cette observation :

Cela dépend du rapport d'autorité des personnes présentes et du degré de formalité du contexte, je trouve qu'en Chine on est en général impoli avec un serveur/euse quand on est client, alors qu'on va l'être avec un médecin.

Je trouve que quand j'habitais dans une famille d'accueil à Hong Kong, la politesse dépendait de la situation. Beaucoup d'activités de la vie quotidienne comme manger au restaurant par exemple sont plutôt considérées comme des activités ordinaires. Par conséquent, on laisse tomber toutes les formules de politesse superflues envers le personnel. Je trouve qu'il ne s'agit pas forcément d'un manque de politesse. C'est juste qu'on part du principe que tout le monde fait son travail sérieusement et que le salaire est la récompense.

Ces exemples montrent que les apprenants se situent entre la deuxième et la troisième étape du processus d'adaptation culturelle. Ils intègrent certains aspects des normes de politesse chinoise et commencent à développer une compréhension de plus en plus approfondie de la société chinoise. Bien qu'ils soient largement influencés par la culture occidentale, ils avancent progressivement vers une meilleure adhésion aux normes de la culture chinoise.

Difficultés

Tous les apprenants ne considèrent pas que l'apprentissage des normes de politesse chinoises soit particulièrement difficile. Pour eux, il s'agit avant tout d'une question de différence culturelle entre la Chine et la France. Cependant, les phénomènes de transfert et de filtrage culturel mentionnés plus haut révèlent un manque de connaissances culturelles et linguistiques chez les apprenants, notamment dans les situations de communication familiales. Notre enquête a permis de mettre en lumière certaines difficultés que les apprenants rencontrent dans l'appropriation des codes de la politesse chinoise, que nous résumons ci-dessous.

Tout d'abord, en raison des profondes différences culturelles entre la Chine et la France, certains apprenants soulignent la difficulté à appréhender la subtilité des codes de la politesse chinoise, voici des témoignages d'apprenants :

Elle [la politesse chinoise] est difficile à appréhender pour moi car les codes de politesse sont entièrement différents de ceux que nous avons en France, et n'ayant que très peu d'amis chinois je n'ai pas l'habitude d'être confrontée à cette politesse si différente de celle que je connais.

À cause des grandes différences culturelles, et de plus, la Chine est très grande. Il pourrait y avoir des différences entre différentes régions.

Cette difficulté est exacerbée par un contexte d'apprentissage non immersif, où les apprenants n'ont pas facilement accès à un enseignement explicite des normes culturelles. Même lorsqu'ils sont en Chine, ils doivent souvent se fier à leur intuition pour comprendre ces codes. Deux apprenants ont d'ailleurs partagé les commentaires suivants :

Il n'y a pas de cours complets sur le sujet et on est confronté à très peu de situations de ce genre en France.

Je pense que la manière de « nommer », de s'adresser à quelqu'un, est assez subtile, je pense que j'ai pu la comprendre en vivant en Chine et en analysant les différents contextes et plus ou moins la classe sociale de la personne en question. Mais cela fait trois ans donc je ne suis plus sûre.

Pour certains, mémoriser les formules de politesse ne pose pas de problème majeur. Cependant, les utiliser efficacement en situation de communication avec des locuteurs natifs chinois s'avère plus délicat, comme l'indique cet apprenant dans le questionnaire :

Je pense qu'elle n'est pas difficile à appréhender, mais qu'elle est difficile à insérer dans notre vie. Par exemple, en étudiant le chinois j'ai appris beaucoup de phrases ou réactions typiquement chinoises dans une conversation, mais quand je parle avec des Chinois, il est difficile pour moi d'insérer ces habitudes dans mon mode d'expression.

Une autre difficulté concerne le sentiment d'identité contradictoire chez les apprenants d'origine chinoise. Nés en France de parents chinois ou d'origine chinoise, certains se trouvent parfois confrontés à un conflit identitaire. Selon Ribert (2009), le sentiment d'identité varie en fonction des situations et des interlocuteurs, et cette variabilité peut être accentuée chez ceux qui possèdent des appartenances culturelles multiples. Les étudiants d'origine chinoise, marqués par une double culture franco-chinoise, peuvent éprouver un sentiment de désorientation lors de leurs échanges avec des Chinois. L'un d'eux a exprimé sa frustration dans le questionnaire :

Je suis d'origine chinoise dans une famille chinoise (avec mes oncles et tantes en France aussi que je vois souvent). Cependant, avec l'éducation reçue en France à

l'école, les choses considérées comme polies peuvent se contredire parfois. Cela va dans les deux sens. J'ai plus tendance à me conformer à la politesse française et je ne sais pas vraiment comment agir selon la politesse chinoise. Par exemple, je vais tendre la joue pour faire la bise par habitude mais certains Chinois trouvent c'est sale. Dans un sens je pense que c'est difficile à appréhender comme nous avons grandi en France.

Introduire les formules de politesse en classe

En tant qu'élément de la compétence sociolinguistique, qui consiste en la connaissance et les habiletés requises pour faire fonctionner la langue dans sa dimension sociétale (Conseil de l'Europe, 2018, p. 143), savoir « utiliser les formes de politesse et tenir compte des règles de politesse » fait partie intégrante des objectifs d'enseignement dès les premiers niveaux. Ainsi, à partir du niveau A1, les formules de politesse peuvent être introduites afin, par exemple, « d'établir un contact social de base en utilisant les formes de politesse les plus élémentaires » (Conseil de l'Europe, 2001, p. 95).

Pour le niveau A2, l'apprenant « peut utiliser des formules de politesse simples et courantes pour s'adresser à quelqu'un ou le saluer » (Conseil de l'Europe, 2001, p. 62). C'est à partir du niveau B1 que l'apprenant devient « conscient des règles de politesse importantes et se conduit de manière appropriée » (Conseil de l'Europe, 2001, p. 95) ; pour cela, l'apprenant de niveau B1 doit être :

[...] conscient des différences les plus significatives entre les coutumes, les usages, les attitudes, les valeurs et les croyances qui prévalent dans la communauté concernée et celles de sa propre communauté. (Conseil de l'Europe 2001, 95)

À leur entrée en deuxième année de licence à l'Inalco, les apprenants sont supposés maîtriser les formules de base. Les deux semestres suivants sont donc consacrés à l'étude de situations de communication plus familières entre locuteurs natifs, avec un accent particulier sur l'usage des codes sociaux et des formules de politesse adaptés à ces contextes.

En collaboration avec une équipe d'enseignants, nous avons conçu une méthode de chinois spécifiquement destinée aux apprenants de deuxième année de licence : *Méthode de chinois troisième niveau* (Arslangul et al. 2019). Cette méthode s'inscrit dans la continuité des deux ouvrages précédemment utilisés en année d'initiation et en première année de licence à l'Inalco. Elle vise à préparer les apprenants à atteindre le niveau B2 du CECRL.

L'une des principales caractéristiques de cette méthode réside dans le fait que chaque leçon est centrée sur un thème socioculturel majeur de la société chinoise contemporaine, comme le soulignent les auteurs dans l'ouvrage :

Toutes les activités langagières proposées dans les leçons sont pratiquées avec cet arrière-plan culturel. L'objectif est d'aider l'apprenant dans la maîtrise des moyens

linguistiques et des connaissances culturelles nécessaires pour interagir en chinois à propos de la société chinoise actuelle. (Arslangul *et al.* 2019, 16)

Les codes de politesse, tels que l'utilisation des termes d'adresse, des formules de remerciement et de salutation, sont intégrés dans les différentes leçons. L'enseignement de ces termes et formules s'accompagne d'activités orales mobilisant à la fois les compétences sociolinguistiques, linguistiques et pragmatiques. Par exemple, dans le cours d'expression orale, nous demandons aux apprenants d'inventer et de jouer des dialogues à chaque séance, mettant en scène deux ou plusieurs personnages appartenant à des statuts sociaux différents. Nous attirons leur attention sur la manière dont les personnages expriment les signes de respect et les codes de politesse, en fonction des relations qu'ils entretiennent. Cet entraînement, évalué régulièrement, permet aux apprenants de pratiquer les codes de politesse dans un contexte de communication imaginaire, mais proche de la réalité. En effet, dans la partie orale de notre méthode, un langage naturel est privilégié :

Nous visons à aider les apprenants à comprendre la langue orale [...], telle qu'elle est utilisée par les locuteurs natifs. Il est aussi de développer les capacités d'expression de l'apprenant pour qu'il puisse s'exprimer de la façon la plus naturelle possible sur des sujets actuels. (Arslangul *et al.* 2019, 18)

Outre ces activités, notre méthode recense l'ensemble des fonctions communicatives que les apprenants doivent maîtriser à l'issue des leçons, telles que « critiquer », « réfuter » ou « ordonner ». Nous insistons systématiquement sur l'importance d'un usage approprié des codes de politesse dans la mise en œuvre de ces fonctions communicatives.

Conclusion

Dans le contexte actuel de mondialisation et d'interconnexion croissante entre les nations, la maîtrise des règles de politesse propres à différentes cultures est plus indispensable que jamais. Comprendre les normes sociales et culturelles qui régissent l'usage de la langue permet aux apprenants d'éviter les malentendus et les erreurs de communication. En tant qu'enseignants, notre rôle ne se limite pas à transmettre une compétence linguistique correcte : il consiste également à guider les apprenants vers un usage socialement approprié de la langue, en fonction des contextes et des interlocuteurs. Cet article vise à sensibiliser les enseignants et les chercheurs à l'importance cruciale de l'enseignement de la politesse dans l'apprentissage du chinois. Par cette contribution, nous espérons participer modestement à la diffusion et à la valorisation de la culture de la politesse en contexte sinophone. Plusieurs pistes de recherche émergent pour prolonger ces réflexions. Parmi elles, la question du sentiment d'identité contradictoire chez les apprenants d'origine chinoise mérite une attention particulière. Bien que cette population représente

une part significative des classes de chinois langue étrangère en France, elle reste encore peu étudiée, tant du point de vue de l'acquisition linguistique que de la construction identitaire. Notre étude n'a pas spécifiquement analysé les représentations de la politesse chinoise chez ces apprenants. Toutefois, les témoignages recueillis suggèrent que ces représentations influencent leur intégration, aussi bien au sein des communautés chinoises que des communautés françaises. Ces observations soulignent la nécessité de poursuivre les recherches afin d'éclairer ces dynamiques complexes et de proposer des pistes pédagogiques adaptées.

Bibliographie

- ARSLANGUL Arnaud, GUO JING, LAMARRE CHRISTINE & LI JING, *Méthode de chinois troisième niveau 中级汉语 (下)*, Paris, L'Asiathèque, 2019.
- BI Jiwan 毕继万, « Limao de wenhua tezheng yanjiu 礼貌”的文化特性研究 » [Étude des caractéristiques culturelles de la politesse], *Shjie Hanyu jiaoxue 世界汉语教学 [Chinese Teaching In The World]*, 1 (1996): p. 51-59.
- BROWN Penelope & LEVINSON STEPHEN C., *Politeness: Some Universals in Language Usage*, Cambridge, Cambridge University Press, 1987.
- CONSEIL DE L'EUROPE, *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer*, Paris, Didier, 2001.
- CONSEIL DE L'EUROPE, *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer. Volume complémentaire avec de nouveaux descripteurs*, Strasbourg, Conseil de l'Europe. Division des Politiques éducatives, Service de l'Éducation, 2018.
- GOFFMAN, Erving, *Interaction Ritual: Essays in Face-to-Face Behavior*, New York, Pantheon Books, 1982.
- GU Yueguo 顾曰国, « Limao, yuyong yu wenhua 礼貌、语用与文化 » [Politesse, pragmatique et culture], *Waiyu jiao yu yanjiu 外语教学与研究 [Foreign Language Teaching and Research]*, 4 (1992): p. 10-17.
- Ji Xinxin 姬辛欣, « Jin shi nian xin meiti yujing xia changyong nüxing chengweiyu yanjiu 近十年新媒体语境下常用女性称谓语研究 » [Étude des appellations féminines couramment utilisées dans le contexte des nouveaux médias au cours des dix dernières années], *Jiamusi zhiye xueyuan xuebao 佳木斯职业学院学报 [Journal of Jiamusi Vocational Institute]*, 3 (2020): p. 161-162.
- KÁDÁR Dániel Z, *Politeness in China*, dans D. KÁDÁR & S. MILLS (éd.), *Politeness in East Asia*, Cambridge: Cambridge University Press, 2011, p. 125-146.
- KERBRAT-ORECCHIONI, Catherine, *La conversation*, Paris, Seuil, 1996.

- KERBRAT-ORECCHIONI, Catherine, *Les interactions verbales II*, Paris, Armand Colin, 1992.
- MENG, Xiaoming, « Malentendus interculturels et enseignement de la politesse verbale concernant les formules de remerciement », *Synergies Chine*, 1 (2001): p. 181-191.
- RIBERT, Evelyne, « À la recherche du “sentiment identitaire” des Français issus de l’immigration », *Revue française de science politique*, 3 (2009): p. 569-592.
- SHAO Jingmin 邵敬敏, « “Meinü” miancheng de zhengyi ji qi shehui yuyanxue diaocha “美女”面称的争议及其社会语言学调查 [La controverse autour de l’appellation directe “měinǚ” (“belle femme”) et son enquête sociolinguistique] », *Yuyan wenzi jingyong 语言文字应用 [Applied Linguistics]*, 4 (2009): p. 71-79.
- TRAVERSO, Véronique, « La politesse et les usages dans les interactions: quelques aspects interculturels », *Langues modernes*, 1 (2000): p. 8-19.
- WANG Jianqin 王建勤, « Kuawenhua yanjiu de xin weidu —— xuexizhe de zhongjie wenhua xingwei xitong 跨文化研究的新纬度 —— 学习者的中介文化行为系统 » [A New Dimension of Cross-cultural Studies-the Learner’s System of Interculture Behavior], *Shijie Hanyu jiaoxue 世界汉语教学 [Chinese Teaching In The World]*, 3 (1995): p. 38-49.
- ZHANG, Linchen, *L’acte de requête dans les méthodes de français langue étrangère et de chinois langue étrangère: une approche comparée*. Thèse de doctorat, Université Grenoble Alpes, 2022.
- ZHANG Zhanyi 张占一, « Shi yi jiaoji wenhua he zhishi wenhua 试议交际文化和知识文化 » [Essai de discussion sur la culture communicationnelle et la culture du savoir], *Yuyan jiaoxue yu yanjiu 语言教学与研究 [Language Teaching and Linguistic Studies]*, 3 (1990): p. 15-32.

Lexique mental bi-plurilingue : les données chronométriques inter-alphabet au service d'une approche unificatrice

Madeleine VOGA

ReSO, Université de Montpellier Paul-Valéry

La représentation langagière dans le lexique mental : symbolique ou distribuée ?

Le lexique mental est défini comme la connaissance des mots de sa langue stockée dans le système langagier. Même si l'hypothèse du stockage a été remise en question, il convient de souligner que la plupart de chercheurs en linguistique considèrent qu'il y a une connaissance relative aux mots conservée *quelque part*, qu'elle contienne des formes des mots (*word-forms*), des fragments de celles-ci ou autre type d'unité et/ou sous-unité.

En effet, si cette connaissance n'était pas stockée, si elle n'avait pas été emmagasinée *quelque part* dans le système langagier, comment aurait-on pu expliquer le fait que l'être humain est capable de reconnaître un mot très rapidement, c'est-à-dire qu'il est capable de distinguer entre un mot existant et un mot non-existant (dans la langue dont il est natif, voire dans sa deuxième ou troisième langue), sans se tromper et en étant sûr de sa réponse ? Comment aurait-on pu expliquer le fait que l'on est capable de corriger un locuteur, natif ou non-natif de notre langue, en lui suggérant le « bon » mot/expression à utiliser ou la forme grammaticalement correcte ?

Il existe donc bien une représentation liée aux mots, même si tous les chercheurs ne sont pas d'accord pour admettre l'existence de représentations (langagières) de nature symbolique. Cependant, comme le souligne Bybee (1995), les schémas abstraits fondés sur les relations entre les mots que proposent les modèles

connexionnistes distribués¹, ne parviennent pas à fournir une explication adéquate de la manière dont fonctionne le sujet. Autrement dit, la critique plus générale des modèles de ce type (sur lesquels se fonde tout le paradigme de l'IA tel que nous le connaissons aujourd'hui) concerne le fait qu'il s'agirait de modèles complètement behavioristes, qui arrivent certes à simuler des données et/ou un comportement « humain », mais qui ne nous renseignent en rien sur les processus en œuvre, ni sur le type d'information qui constitue le substrat sur lequel s'appliquent ces processus.

Ces deux questions (le type d'information ainsi que les processus et représentations impliqués dans le traitement langagier) ont fait l'objet de nombre de travaux, dans leur grande majorité dans un cadre théorique compatible avec l'existence d'un lexique mental, c'est-à-dire compatible avec l'idée que la connaissance intériorisée des mots de sa (ou ses) langue(s) est inscrite en mémoire, et plus précisément dans la Mémoire à Long Terme (MLT²).

L'approche dans laquelle s'inscrit la présente contribution est fondée sur l'existence de représentations langagières formées durant l'acquisition et l'apprentissage de la langue ou des langues que maîtrise le sujet. Il s'agit ici de la conception 'classique' de l'étude des représentations mentales liées au langage. Ces représentations sont définies comme locales, symboliques et discrètes (ex. McClelland et Rumelhart, 1981 ; Grainger et Jacobs, 1996), dans le sens où une représentation de nature orthographique sera différente par rapport à une représentation de nature morphologique. L'enjeu donc consiste à étudier ces représentations, leur nature, leur organisation, leurs caractéristiques et leurs rapports. Une telle approche est compatible, en tout cas sur le plan théorique, avec d'éventuelles différences provenant des caractéristiques particulières des langues étudiées.

Si l'on se réfère au lexique bi-plurilingue, se pose également la question des transferts entre les langues. Des travaux expérimentaux que nous présenterons

1. L'idée selon laquelle le langage repose sur des représentations de type symbolique n'est pas acceptée par certains chercheurs, notamment ceux qui se situent dans une perspective connexionniste complètement distribuée, c'est-à-dire sans aucun recours à des connaissances explicites. Ces modèles s'inspirent, de près ou de loin, du modèle « Associateur de Patterns » (*Pattern Associator Model*) de RUMELHART & MCCLELLAND (1986) conçu pour imiter l'acquisition du passé en anglais. Ce modèle connexionniste (réseau de neurones) a appris à associer des caractéristiques phonologiques du thème du verbe (ex. *walk*) aux caractéristiques phonologiques de la forme du passé (*walked*). Puisqu'un réseau de neurones sans connaissance stockée peut « apprendre » le passé anglais à partir de traits orthographiques et phonologiques, la représentation symbolique devient superflue. Cette approche a conduit à un grand nombre de travaux et de modèles, par ex. PLAUT, MCCLELLAND, SEIDENBERG & PATTERSON (1996), simulant des résultats de lecture de mots en anglais ou celui de SEIDENBERG & MCCLELLAND (1989) qui décrit l'acquisition du passé comme un continuum quasi régulier de correspondances entre thèmes (stems) et formes du passé. Tous ces travaux récuse l'existence de représentations langagières, misant tout sur l'idée qu'un mot est représenté par la configuration d'activation à travers un ensemble de traits (représentation distribuée). C'est ainsi que la morphologie du verbe anglais est devenue le terrain de prédilection de l'affrontement entre les explications connexionnistes distribuées et les explications symboliques, par ex., MCCLELLAND & PATTERSON (2002) et RAMSCAR (2002) *versus* PINKER & ULLMAN (2002).

2. La localisation exacte et le rôle des différents types de mémoire ne font pas l'objet de cette contribution.

brèvement ci-dessous, démontrent que les transferts entre les langues du bilingue reposent sur l'existence de représentations connectrices (*linking representations*), autrement dit des connexions inter-langues. La question qui est abordée dans ces travaux concerne la spécification des représentations qui opèrent le lien entre les deux « parties » du lexique bilingue, sans que cela signifie que ces dernières soient séparées. Comme nous le verrons par la suite, ces deux « parties », par exemple « la partie grecque » et la « partie française » du lexique bilingue, se trouvent dans un rapport étroit et une interaction permanente durant le traitement. Parmi les travaux qui se donnent comme objectif la spécification des représentations connectrices (entre les deux langues), plusieurs prennent comme point de départ l'approche théorique de Bybee (1985 ; 1988) qui décrit le lexique monolingue comme une série de paradigmes lexicaux (*clusters*), formés par la base et ses dérivations. Cette organisation est supposée dépasser (*transcend*) les langues, elle s'applique donc à la fois au lexique monolingue et au lexiques bilingue et bi-plurilingue. Elle est fondée sur l'idée que les représentations qui assurent la connexion au sein du lexique mental seraient de nature morphologique. Les connexions se créent et se développent au fur et à mesure de l'acquisition (d'une L1 et/ou d'une L2) grâce à l'usage ; c'est pour cela que l'on se réfère à une approche *émergente* (Ellis 2006) de la compétence lexicale et morphologique chez le locuteur (L1 et/ou L2). L'approche théorique que nous esquisserons dans cet exposé s'inscrit dans l'approche dite « intégrative », c'est-à-dire qui vise à combiner et éventuellement concilier l'approche *bottom-up* et *top-down*. Vu que nous nous intéressons à une combinaison de langues, le grec et le français, qui ne s'écrivent pas avec le même alphabet, les facteurs liés aux hauts niveaux de traitement (*core factors*, cf. Mulder, Dijkstra & Baayen, 2015) entrent en jeu de manière plus directe et facilement observable, en tout cas par rapport à une situation expérimentale dans laquelle il n'y a pas de différence d'alphabet (ex. anglais-français). Comment en pourrait-il être autrement, lorsque le lexique bilingue associe, selon l'expression de Grosjean (2001) *one mind two languages*, un même contenu sémantique (et/ou conceptuel¹) à deux formes totalement différentes du point de vue formel ?

L'effet cognat dans l'amorçage inter-langues : brève revue d'études intra- et inter-alphabet

Un grand nombre de recherches publiées depuis les 30 dernières années examinent le traitement (accès lexical, *lexical access*) de différents types de mots, dans différents

1. Comme le fait déjà remarquer Grosjean (1998) l'usage des termes *sémantique* et *conceptuel*, de ce qui distingue ces deux termes et de ce qui les rapproche, n'est pas toujours clair dans la bibliographie relative au bilinguisme, qui, autour des années 1998, correspondait à un domaine nouveau. Si l'on considère que le terme « conceptuel » recouvre des traits et des aspects relevant de l'extralinguistique, on peut se demander si cette surface conceptuelle correspond à la L1 ou à la L2. À mon avis, ce problème terminologique persiste jusqu'à aujourd'hui.

types de langues et leurs combinaisons, avec une prépondérance, comme nous pouvons aisément l'imaginer, de couples de langues qui incluent l'anglais. Pour ce qui concerne le lexique bilingue, de protocoles chronométriques et en particulier la technique de l'amorçage masqué (*masked priming*) ont été employés afin d'examiner le traitement de différents types de cognats (ce type de protocole expérimental fera l'objet d'une présentation relativement détaillée plus loin).

Dans un tel protocole, la condition qui comporte des mots cognat (ex. *palacio-palace* en esp.-anglais) sera plus rapide à traiter, c'est-à-dire à reconnaître, que la condition de contrôle non-relié¹. Autrement dit, l'amorce (en espagnol) *palacio* parvient à introduire un bénéfice, de l'ordre de 50-60 ms en général, sur la reconnaissance de la cible (en anglais) *palace*. Il s'agit donc d'un transfert d'activation d'une langue à l'autre, qui apparaît dès les tout premiers stades de la reconnaissance lexicale, puisque la technique de l'amorçage masqué étudie précisément ces stades précoces de l'accès lexical. Ce type d'effet cognat est bien documenté dans une variété de langues, l'on parle alors d'amorçage inter-langues (*cross-language priming*), que ce soit pour des couples de langues qui s'écrivent avec le même alphabet (conditions intra-alphabet), cf. les études classiques de Cristofanini, Kirsner & Milech (1986), avec des cognats espagnol-anglais ; hollandais-anglais, de Groot & Nas (1991), ou bien, pour des couples de langues inter-alphabet, ces dernières ayant connu un développement plus tardif, ex. Gollan, Forster & Frost 1997, en hébreu-anglais ; Jiang & Forster 2001, en chinois-anglais ; Voga & Grainger 2007, en grec-français ; Iakovleva, Piasecki & Dijkstra 2015, en russe-anglais ; Miwa, Dijkstra, Bolger & Baayen 2014, en japonais-anglais. Plusieurs variantes de ce protocole, *i.e.* plusieurs types d'amorçage, ont été employés, cf. par ex. Dijkstra, Van Jaarsveld, & Ten Brinke (1998), Dufour & Kroll (1995), Lavour & Font (1998), Sánchez-Casas & García-Albea (2005), y compris des protocoles dans lesquels des cognats de trois langues étaient présentés aux participants, par ex. dans l'étude de Lemhöfer, Dijkstra & Michel (2004) en anglais-allemand-hollandais.

Approches *bottom-up* (perception) et *top-down* (production) et leurs implications pour l'étude du bilinguisme inter-alphabet

Dans le cadre des modèles issus de l'approche *bottom-up*, où la primauté est donnée à des facteurs sous-lexicaux et formels (notamment orthographiques), les cognats doivent partager une grande part de leur forme orthographique ou, à défaut, phonologique, *e.g.* *sport-sport* en français-anglais (Dijkstra, Grainger et Van Heuven 1999). Ce type de modèle se focalise sur des effets liés à la perception du langage écrit, de fait caractérisée par une visée *bottom-up*, c'est-à-dire émanant des caractéristiques visuelles des lettres et se dirigeant vers le niveau orthographique et, *in fine*, la représentation du mot. C'est le positionnement de l'un des

1. Une condition de contrôle non-relié est une amorce qui ne partage aucune similarité sur les plans orthographique, phonologique, morphologique et sémantique avec la cible correspondante.

modèles les plus influents de l'accès lexical, le *Bilingual Interactive Activation Model* (BIA, Dijkstra, Grainger et van Heuven 1999). Initialement, ce modèle était une extension du modèle d'activation interactive de la reconnaissance visuelle du mot, à l'origine proposé par McClelland and Rumelhart (1981), qui, en tant que tel, n'avait pas codé de représentation sémantique. La sémantique a été introduite dans une version ultérieure de ce modèle, BIA+ (Dijkstra & Van Heuven 2002), en même temps que l'information phonologique, dans le but de le rendre capable de capturer des effets cognat inter-alphabet (*e.g.* μετρό /me'tro/ - *méto*). Cette nouvelle version (BIA+) du modèle comporte aussi un schéma de tâche, destiné à capturer le fait que l'identification du mot bilingue doit aussi refléter le contexte de la tâche défini par les conditions expérimentales, par ex. s'il s'agit d'une tâche monolingue ou bilingue.

Il faudrait faire remarquer ici que la bibliographie, largement anglo-centrée, surtout au début du développement de cette ligne de recherche, est marquée par la tentation de limiter l'attribution du statut de cognats à un certain type de mots, c'est-à-dire aux mots qui partagent une similarité orthographique (et phonologique) parfaite, comme *sport-sport*. Dijkstra, Grainger et Van Heuven (1999) par exemple, proposent de distinguer entre cognats de similarité ortho-phonologique maximale et ceux de similarité ortho-phonologique réduite, de type κέντρο /'kedro/ - *centre*, ces derniers portant l'étiquette de « semi-cognates¹ ». On peut faire remarquer que, pour une approche qui vise à décrire du point de vue cognitif tout ce qui a trait à la manière dont les locuteurs identifient les mots de leurs deux langues, il apparaît quelque peu négligeant de catégoriser les cognats uniquement sur la base de leur similarité orthographique-visuelle, ou leur similarité phonologique, s'il s'agit de cognats s'écrivant avec des alphabets différents. Alors même que des études inter-alphabet qui démontrent l'existence d'effets d'amorçage tout à fait équivalents à ceux d'études intra-alphabet, sont publiées dès 1997 (Gollan, Forster & Frost 1997, en hébreu-anglais) et au début des années deux mille (Jiang & Forster, 2001, en chinois-anglais). Cependant, et sans que ceci soit une justification pour cette négligence, il convient de reconnaître qu'il a fallu une grande quantité de travaux, venant d'une variété de langues, pour qu'une approche plus « inclusive » puisse émerger. Par ailleurs, il faudrait souligner que les protocoles expérimentaux d'amorçage, surtout masqué, peuvent présenter une grande dépendance aux facteurs orthographiques, comme le reconnaissent les chercheurs qui ont inventé ces techniques (ex. Forster, Mohan et Hector 2003), d'où de nombreuses précautions méthodologiques prises dans la planification des études expérimentales.

1. On retrouve exactement ce même type de raisonnement auprès de certaines théorisations qui valident l'hypothèse d'une difficulté supposée accrue en raison du changement d'alphabet. Comme nous le montrerons dans la suite de cet exposé, ceci correspond à une vision superficielle du langage, qui accorde une nette priorité aux facteurs formels (orthographe, phonologie) au détriment des facteurs morphologiques et sémantiques-conceptuels.

Nous pouvons compléter ce bref historique concernant l'approche *bottom-up* en précisant que la situation tend à nettement s'améliorer à partir de 2012-2013, lorsque plusieurs travaux impliquant le chinois et le japonais voient le jour (ex. Miwa, Dijkstra, Bolger et Baayen 2014, en japonais-anglais). À partir du moment où on accepte, preuves expérimentales à l'appui, qu'il puisse y avoir des « vrais » cognats japonais-anglais ou chinois-anglais, malgré la différence considérable qui sépare les deux langues sur le plan de la graphie, alors cela veut dire que la différence au niveau visuel-orthographique n'est pas aussi déterminante qu'initialement supposé. Cette acceptation a ouvert la voie à un courant de recherche portant sur l'étude de paires de langues « éloignées », parmi lesquelles nous pouvons ranger le bilinguisme grec-français (ex. Voga & Grainger 2007) qui a conduit, entre autres, à « l'explication phonologique de l'effet cognat », applicable en particulier à des couples de langues comme le japonais-anglais (Nakayama, Sears, Hino & Lupker 2012 ; Nakayama, Verdonchot, Sears & Lupker 2014) ou le chinois-anglais (Wang, Wang & Malins 2017 ; Wang & Li 2025). Indépendamment de la pertinence d'une telle répartition, étant donné que le lexique mental bilingue ne peut pas être réduit à des similarités en termes de graphèmes et/ou phonèmes, et que la dichotomie inter- *versus* intra-alphabet comporte une tendance à la surgénéralisation, force est de constater que cette ouverture vers des langues « autres » que les langues européennes a été bénéfique au domaine du bilinguisme et du plurilinguisme en général.

Aux antipodes de l'approche *bottom-up*, se trouve l'approche issue de la production du langage, qualifiée de *top-down* (e.g. RHM, *Revised Hierarchical Model*, Kroll et Stewart 1994 ; pour la dernière version, cf. Kroll, van Hell, Tokowicz & Green 2010), qui attribue une place prépondérante au niveau conceptuel. La définition de la relation cognat devient par conséquent moins stricte sur le plan formel, il est donc facilement admis que des mots avec une similarité formelle réduite (e.g. *rain* — *regen*, dans l'étude *princeps* de Groot et Nas 1991) puissent induire des effets équivalents à ceux des cognats identiques sur le plan de la forme.

Il serait utile ici de brièvement rappeler le fonctionnement du modèle RHM, qui demeure l'un de modèles de compétence bilingue les plus influents. Il a été aussi le premier modèle à prendre en compte des données de production bilingue et pour cette raison a été influencé par les données en faveur de l'accès lexical sélectif, position qui est aujourd'hui abandonnée. Ce modèle opère une distinction entre le niveau lexical et le niveau conceptuel. Au niveau lexical, les deux lexiques sont distingués, un pour les mots de la langue « maternelle » (L1) et un autre pour les mots de la langue seconde (L2). Ces deux lexiques sont connectés dans un système conceptuel commun, qui contient le sens des mots. Ce modèle pose que les liens, aussi bien lexicaux que conceptuels, sont bidirectionnels, mais qu'ils diffèrent en force-prégnance (*strength*) : le lien lexical de la L2 vers la L1 est supposé être plus fort que le lien lexical de la L1 vers la L2, étant donné que les mots de la L2 ont été initialement associés à la L1. De ce point de vue, le

modèle est hiérarchique. De même, le lien de la L1 à la mémoire conceptuelle est supposé être plus fort que celui de la L2 vers la mémoire conceptuelle (Kroll et Stewart, 1994 : 158). Il doit être souligné ici que ce modèle a été créé afin de faire des prédictions pour les effets de traduction de la L1 vers la L2 et inversement, et ceci spécialement dans le cadre de protocoles de production. Par conséquent, les aspects sémantiques et conceptuels revêtent une importance particulière et ceci est l'une des raisons pour lesquelles ses auteurs assument des connexions de force différente entre le niveau lexical et conceptuel pour les deux langues.

Si l'on se place du point de vue des implications des modèles *bottom-up* vs *top-down*, nous pouvons souligner que la « hiérarchisation des besoins », qu'opère le BIA, autrement dit, le fait que la primauté soit attribuée aux facteurs formels, induit un vrai déséquilibre et ne parvient pas à une prise en compte suffisante des aspects liés à la sémantique et/ou la morphologie. Celle-ci est la critique la plus importante que l'on pourrait adresser à ce type de modèle, en tout cas du point de vue morphologique, qui est celui qui nous intéresse ici. Ce type de lacune est relativement compréhensible, dans la mesure où le BIA et le BIA+ constituent des modèles de perception et non pas de production (comme le RHM). On aura compris que ces deux approches, celle du BIA+ et celle du RHM, l'une fondée sur la perception, l'autre sur la production, quoique complémentaires, poursuivent des chemins parallèles.

Pour ce qui concerne le modèle RHM, nous souhaitons souligner que l'un de ses mérites est qu'il a essayé de répondre à la question des interactions entre les différentes composantes du lexique bilingue à un niveau central (*core level*), c'est-à-dire au niveau sémantique — conceptuel, à travers l'investigation des connexions entre le lexique L1 et le niveau conceptuel, celui de la L2 et le niveau conceptuel, etc. Ceci peut constituer le point de départ de l'interrogation relative à la connexion, possiblement de nature morphologique, entre les deux lexiques du bilingue.

Les transferts d'activation au sein du lexique mental et les similarités entre les langues du bilingue

Cette question, relative aux transferts ou aux représentations connectrices entre les langues du sujet bilingue et plurilingue, ne peut pas être considérée en dehors des similarités entre langues. Reste à voir ce que nous entendons par ces « similarités ». Comme nous l'avons vu plus haut, l'étude des cognats a été fortement marquée, pendant longtemps, par l'influence de facteurs orthographiques, la raison pour cela étant, entre autres, que les recherches sur l'accès lexical bilingue concernaient, en tout cas au départ, des protocoles intra-alphabet. Pour les langues qui s'écrivent avec des alphabets différents, cette similarité sur le plan de la forme est uniquement phonologique, ex. *παλάτι*-*palais* en grec-français. Il serait donc possible que les effets d'amorçage observés pour les cognats soient en

réalité de simples effets d'amorçage formel, plutôt que des effets de traduction, ce qui voudrait dire que la représentation sémantique (ou conceptuelle, dans un modèle comme le RHM) ne serait pas contactée. Cette hypothèse doit être exclue, dans sa version dure, il demeurerait néanmoins, surtout avant le développement d'études inter-alphabet, que la participation purement formelle n'était pas facile à évaluer dans des protocoles qui employaient un couple de langues qui s'écrivent avec le même alphabet (*i. e.*, sous des conditions intra-alphabet). Une difficulté supplémentaire venant du fait que le degré de similarité phonologique que partagent les cognats n'est pas forcément corrélé avec leur similarité visuelle, ni avec la distance qui sépare les deux langues : deux cognats anglais — hollandais, ex. *brother* — *broer* « frère » (de Groot et Nas 1991) peuvent avoir moins de similarité phonologique que deux cognats anglais — hébreu comme *pyramid* — *pyramida* (Gollan *et al.* 1997) qui n'ont pourtant aucune relation orthographique (alphabet latin et alphabet sémitique). Le couple de langues grec-français (ou français-grec, dans l'autre direction d'amorçage, L2 vers L1, que nous n'aborderons pas ici) peut donc nous aider à distinguer, dans un premier temps, entre l'influence des facteurs formels et sémantiques, pour qu'ensuite différents types de combinaison de ces deux facteurs puissent être étudiés.

Par ailleurs, comme nous le verrons par la suite de cet exposé, les similarités entre langues (*language similarities*) devraient, à notre sens, inclure les relations paradigmatiques entre les mots, en particulier les relations qui peuvent se créer, au fur et à mesure du développement bi-plurilingue, entre une partie du signifiant et/ou du signifié de la L1 et une partie du signifiant et/ou du signifié de la L2. Ce point sera abordé dans la section consacrée aux effets morphologiques.

L'amorçage masqué

La méthodologie de l'amorçage emploie l'effet dit « de répétition », *i. e.* le fait qu'un même stimulus, présenté deux fois, est reconnu plus rapidement par le sujet lors de sa deuxième occurrence. À partir de cette constatation, les expériences basées sur l'effet d'amorçage visent à vérifier si des temps de réaction (TR) plus rapides sont également obtenus pour des stimuli qui, bien que différents par rapport au stimulus « répétition » (du point de vue orthographique, sémantique, morphologique), seraient reliés à celui-ci en vertu d'une similarité formelle, sémantique ou d'une relation morphologique (*i. e.* à la fois formelle et sémantique). En d'autres termes, la question ici concerne l'impact qu'auront, en termes de TR, les différentes relations entre les mots. C'est pour cela que parmi les différentes conditions qui sont planifiées dans une expérience, il y a toujours des conditions de contrôle (cf. Tableau 1, pour un exemple simple). Ce contrôle est, la plupart de fois, une condition neutre, que nous appelons la condition « non-reliée », qui représente l'absence de relation. Il peut cependant y avoir des contrôles plus ciblés, comme par ex. des contrôles phonologiques sont employés dans l'étude de Voga

& Grainger (2007), dans le but de capter la contribution clairement attribuable à la similarité phonologique entre amorce et cible.

Tabl. 1 — Illustration d'un exemple de protocole d'amorçage masqué en français. Les effets d'amorçage (fictifs) sont estimés par rapport au contrôle non-relié

Cible	Conditions d'amorçage et amorces			Effets nets d'amorçage	
	Répétition	Morphol. (Morph.)	Non-relié (NR)	NR-Rep.	NR-Morph.
Monter	monter 500 ms	montais 530 ms	poster 600 ms	100 ms*	70 ms*

Cependant, le fait de « répéter » les stimuli, même s'ils ne sont pas répétés à l'identique (la cible *monter* n'est pas identique à l'amorce *montais*), peut induire des stratégies de la part du sujet, qui aura fini par comprendre le type de rapport que partagent l'amorce et la cible. C'est pour cela que les chercheurs emploient, depuis les années 1990, l'amorçage masqué, dans lequel la première apparition du mot, qui est appelée « amorce » (*prime* en anglais) n'est pas perçue de manière consciente par le sujet. Le tableau 1 donne un exemple simple de ce type. Pour la (même) cible *monter*, plusieurs conditions d'amorçage sont créées : une condition répétition (*monter*), qui représente la condition qui induira un maximum de facilitation, une condition morphologique, ici une flexion (*montais*), qui est celle dont nous cherchons à estimer l'effet précis, et une condition de contrôle non-relié, qui représente la condition neutre (*poster*). La cible, différente à chaque essai, sera précédée par une amorce qui apparaîtra à l'écran pour une durée déterminée (SOA, *Stimulus Onset Asynchrony*) et extrêmement courte, de l'ordre de 48 ms. Cette durée ne permet pas l'identification consciente de la cible, mais suffit pour que la cible soit traitée par le système langagier du participant normo-lecteur.

Comme on vient de le voir, en fonction des hypothèses de l'expérience, l'amorce peut partager différents types de relation avec la cible. Si l'amorce correspond exactement à la cible (condition « répétition »), les TR (temps de réaction) pour l'identification de la cible et la réponse du sujet (réponse motrice incluse) seront plus courts, comparés aux autres conditions, ici 500 ms. Inversement, si l'amorce n'a aucun lien avec la cible (condition non-reliée), aucune facilitation n'est attendue sur les TR (ici, 600 ms). Ces deux conditions (répétition et non-reliée) nous servent de base afin d'estimer les effets de la ou des conditions « critiques ». Dans l'exemple de la table 1, nous observons un effet de répétition de 100 ms, par rapport au contrôle non-relié. Cet effet représente la facilitation maximale. Nous observons aussi un effet morphologique de 70 ms, toujours par rapport au contrôle non-relié. Cet effet représente l'effet d'amorçage morphologique.

Au niveau de l'interprétation, la facilitation de 70 ms par rapport au contrôle non relié (Table 1) est interprétée comme preuve d'une organisation morpholo-

gique du lexique, à condition d'être significative. L'idée sous-jacente à cette interprétation est que *montais* n'aurait pas pu activer (ou amorcer) la cible *monter*, si ces deux unités n'étaient pas interconnectées, en vertu de leur parenté morphologique, si elles n'étaient pas reliées par une connexion, au sein du lexique mental. Cette activation ne peut pas faire l'objet d'une stratégie de la part du sujet, étant donné qu'il n'a pas perçu consciemment l'amorce. Cette activation est donc automatique (Forster, Mohan & Hector, 2003), et concerne des processus automatiques et irrépressibles chez le lecteur expert, durant les tout premiers stades de l'identification du mot, auxquels la bibliographie réfère par le terme accès lexical (*lexical access*). L'activation de la cible par l'amorce est prise comme preuve du fait que ces deux représentations sont interconnectées, ou appartiennent au même *cluster*, pour exprimer cette idée avec les mêmes termes que Bybee (1995 ; 2007), de telle manière que l'activation de la représentation de l'une entraîne automatiquement, et dans des délais extrêmement courts, l'activation de l'autre, en l'absence de toute stratégie prédictive ou contrôle conscient de la part du sujet-participant.

Ainsi, l'amorçage masqué est devenu la technique privilégiée pour observer les processus automatiques et irrépressibles (chez le lecteur expert) qui ont lieu pendant l'identification lexicale ou accès lexical, et qui reflètent le traitement cognitif dont a fait l'objet le stimulus. Nous sommes donc dans une perspective qui se situe clairement à l'opposé de ce que nous avons décrit plus haut (note de bas de page 1), concernant les modèles à représentations totalement distribuées, puisque ce qui nous intéresse ici n'est point le résultat du traitement (« est-ce que le modèle informatique/réseau neuronal a appris ou pas ? »), mais plutôt le chemin, la manière dont tel ou tel stimulus (ou groupe de stimuli présentant les mêmes caractéristiques) a été traité par le *vrai* sujet parlant.

Illustration du paradigme d'amorçage masqué inter-langue inter-alphabet (grec-français)

Dans cette section, nous présenterons brièvement l'étude de Voga & Grainger (2007), l'une des premières avec des données d'amorçage masqué inter-langue et inter-alphabet, étude qui d'une part nous servira de « base » pour décrire plusieurs ensembles de données issues de la même méthodologie, et qui d'autre part, a contribué à démontrer l'intérêt de combinaisons de langues inter-alphabet. Cette étude démontre, *via* trois expériences d'amorçage masqué auprès de sujets bilingues grec-français, que l'effet cognat est sensible à la similarité phonologique entre les mots des deux langues (en l'absence de similarité orthographique, dans une situation inter-alphabet) et est donc dû à un transfert d'une langue à l'autre au niveau de la forme (phonique) et du sens. L'expérience 2 de Voga & Grainger (2007, 942), que nous présenterons ici en quelque détail, a été menée auprès de 30 sujets grecs, apprenants avancés du français et ayant séjourné au moins un an en

France, notamment pour les études. Ils ont tous le grec comme langue première (maternelle) et disposent tous de diplômes FLE.

Au niveau de la planification expérimentale, toutes les amorces étaient en grec (L1 des participants) et toutes les cibles en français (L2), il s'agit donc d'un protocole d'amorçage masqué inter-langue, dans la direction L1 vers L2, dans lequel la durée de présentation de l'amorce est de 48 millisecondes (ms). Dans le cadre de ce type de protocole, réalisé grâce au programme DMDX (Forster & Forster 2003), la perception de l'amorce en grec, présentée pendant quelques millisecondes (SOA de 48 ms) influe sur l'identification de la cible, sans que les participants aient eu le temps de consciemment identifier¹ l'amorce. La tâche qui est demandée aux participants est une tâche de décision lexicale, c'est-à-dire ils doivent « décider le plus correctement et le plus rapidement possible si le mot qui apparaît sur l'écran de l'ordinateur est un mot — ou pas ». Les matériaux (stimuli) pour cette expérience étaient :

- a. 30 cognats caractérisés par une similarité phonologique maximum (92,6 %) entre le grec et le français, comme *diva* ou *piano*, *métro*, qui donnent $\nu\tau\acute{\iota}\beta\alpha$ /'diva/, $\pi\acute{\iota}\alpha\nu\omicron$ /'piano/ ou encore $\mu\epsilon\tau\rho\acute{o}$ /me'trol/ en grec ;
- b. 30 cognats dont la similarité phonologique était moins importante (55,2 %), comme $\mu\omicron\upsilon\sigma\epsilon\acute{\iota}\omicron$ /mou'sio/ - *musée*, $\kappa\acute{\epsilon}\nu\tau\rho\omicron$ /'kedrol/ - *centre* ou encore ;
- c. enfin, 30 mots non-cognats, ex. $\kappa\lambda\acute{\alpha}\mu\alpha$ /'klama/ - *pleur*, qui ne partagent aucune similarité phonologique, ni orthographique ou morphologique-étymologique².

Pour cette expérience, il y avait trois conditions d'amorçage, une amorce traduction et deux conditions de contrôle : une condition de contrôle non-relié, ainsi qu'un contrôle phonologique, c'est-à-dire un contrôle sans aucune relation morphologique avec la cible, qui partage avec cette dernière autant de phonèmes

1. La procédure, commune à toutes les expériences présentées ici, est la suivante : Le protocole est conduit sur PC avec le logiciel DMDX (FORSTER & FORSTER 2003). Chaque essai consiste en trois événements visuels : le premier est un « masque » proactif formé de 10 caractères dièses (#) d'une durée d'apparition de 500 millisecondes. Le masque est immédiatement suivi par l'amorce avec une durée d'exposition (SOA) de 48 ms. L'amorce est immédiatement suivie par la cible qui reste à l'écran jusqu'à ce que le sujet réponde. En cas d'absence de réponse du sujet, le programme passe à l'item suivant au bout de 4 secondes. L'espace inter-stimulus est de 500 ms. Tous les stimuli apparaissaient au milieu de l'écran en lettres minuscules et en alphabet grec. Les minuscules ont été préférées aux majuscules, d'une part, parce que les majuscules ne sont pas usuellement utilisées en grec (en dehors de l'usage standard, identique au français, ex. noms propres), et d'autre part, afin de préserver l'accent tonique placé au-dessus de la lettre appropriée, les majuscules s'écrivant sans accent tonique. Pour éviter qu'il y ait un recouvrement orthographique entre l'amorce et la cible, les amorces sont présentées en caractères plus petits que les cibles (Times New Roman de taille 12 et 16 respectivement). Les sujets étaient assis à une distance d'environ 50 centimètres de l'écran de l'ordinateur. Après 12 essais de familiarisation, les 180 items expérimentaux étaient présentés en une série continue. Nous avons donné aux sujets la consigne de faire des décisions lexicales sur les mots présentés à l'écran aussi rapidement et aussi correctement que possible.

2. Dans cette expérience de décision lexicale, il y a également 90 non-mots (ou pseudo-mots), afin que les participants puissent choisir entre la réponse « Oui, c'est un mot » et « Non, ce n'est pas un mot ».

que la cible partage avec son cognat. Par exemple, pour la cible (de similarité phonologique réduite) *chimie* dont le cognat grec est χημεία /him'ia/, les phonèmes partagés sont les trois phonèmes soulignés (him'ia) que l'on retrouve dans le contrôle phonologique /sim'ia/ « pointPL » ; pour un stimulus de similarité phonologique maximum comme μετρό — *méto*, le contrôle phonologique est μετρώ /me'tro/ « compter », ainsi la forme phonologique du mot est identique, position de l'accent compris¹. Cette condition de contrôle phonologique vise à évaluer la participation exacte du facteur phonologique, puisque l'effet cognat (*i.e.*, l'effet de traduction) pourrait être différent en fonction du contrôle utilisé, vu que la similarité entre le mot grec et français est uniquement phonologique. Les résultats de cette expérience montrent que les amorces des deux niveaux de similarité phonologique engendrent des effets (statistiquement) équivalents, lorsque les effets sont estimés par rapport au contrôle phonologique : 38 ms pour les cognats de similarité phonologique maximale, comme μετρό-*méto* et 46 ms pour ceux de similarité minimale, comme χημεία-*chimie*. Lorsque les effets sont évalués par rapport à un contrôle non-relié, les effets sont également robustes (et statistiquement significatifs) pour les deux types de matériaux, cognats de similarité phonologique maximale (55 ms d'effet) et minimale (45 ms). Enfin, la condition de non-cognat, ex. λάθος-*erreur* parvient à induire un effet significatif de 36 ms, par rapport au contrôle phonologique (et 23 ms par rapport au contrôle non-relié, cf. Table de résultats, Voga & Grainger 2007, 943). Le résultat critique de cette étude concerne le pattern différent des effets selon le type de contrôle utilisé. Quand les effets sont mesurés relativement à la classique condition non reliée, le degré de similarité formelle des cognats (et non-cognats) affecte directement l'amorçage observé. Quand la condition de contrôle est appariée en termes de similarité formelle à la condition de traduction, le degré de similarité phonologique n'affecte plus l'amorçage, et les deux catégories de cognats se comportent presque de la même manière. Le fait qu'un effet non-cognat est aussi obtenu, entre l'amorce grecque, ex. κλάμα /'klama/, « pleur » et son mot-cible correspondant en français (*pleur*), alors que ce dernier n'est pas systématiquement trouvé dans les études intra-alphabet, est interprété dans la même direction : dans l'amorçage inter-langue grec-français, il y aurait autre chose que de simples effets de forme, et plus précisément une participation du sens. Le sens du mot grec est activé, durant ces premiers stades d'identification lexicale, et c'est cette activation du sens, accompagnée de l'information phonologique, disponible dès les premiers stades de la reconnaissance lexicale bilingue, qui induit les effets d'amorçage significatifs. Ces effets² sont robustes (entre 46 et 36ms par rapport au contrôle phonologique)

1. Seule le graphème final, oméga *versus* omikron, distingue ces deux formes, l'emprunt μετρό « méto » et le verbe μετρώ « compter ». Sur le plan phonologique, ces deux formes sont donc identiques.

2. Tous les effets auxquels nous faisons référence ici sont statistiquement significatifs. Nous remercions le lecteur de se rapporter à la publication citée pour l'analyse statistique complète que nous ne pouvons fournir ici, faute de place.

et récurrents, puisqu'ils se manifestent pour les trois types de stimuli, ceux de similarité phonologique maximale, minimale, ainsi que dans le cas où il n'y a aucune similarité formelle (non-cognats). Ces effets sont donc indépendants du niveau de similarité formelle, puisqu'ils sont trouvés même pour les non-cognats, dans cette expérience. Ils démontrent que l'effet d'amorçage induit par les cognats reflète la contribution de deux mécanismes distincts : l'un sensible à la similarité sémantique entre amorce et cible, et l'autre sensible à la similarité formelle.

L'impact de différents types de rapports qu'entretiennent les mots : le rôle de la morphologie

Du point de vue de l'acquisition et du traitement, la morphologie lexicale est sans doute l'un des domaines les moins explorés par la linguistique traditionnelle pour des raisons à la fois théoriques et méthodologiques. Ces raisons seraient liées, au moins partiellement, à une vision où la morphologie lexicale correspond à la capacité de l'apprenant à combiner de manière réussie des morphèmes lexicaux et grammaticaux (*word-formation rules*, e.g. Aronoff, 1976 ; Scalise, 1994) et où les mots complexes résultent de cette combinaison. Dans un cadre de ce type, la question principale consiste à déterminer si et quand un apprenant a effectivement recours à des mécanismes morphologiques pour reconnaître, comprendre ou produire des mots complexes. Il semble y avoir un consensus autour de l'idée que, durant les premiers stades de l'acquisition, l'apprenant de la L2 traite le mot de manière holistique, comme une unité non analysable. Par la suite, et après une exposition suffisamment prolongée à la L2, il devient capable de reconnaître sa structure et éventuellement produire le mot. L'émergence d'une 'stratégie morphologique' est toutefois difficile à déterminer : le fait que l'apprenant produise un mot complexe ne signifie pas nécessairement qu'il possède une compétence dérivationnelle, et ne nous renseigne en rien sur les processus mis en œuvre pour la perception et la production de telle ou telle forme morphologiquement complexe. La forme produite peut très bien être directement récupérée dans la mémoire à long terme (MLT), sans qu'il n'y ait aucune compétence morphologique en L2. C'est d'ailleurs ce qu'envisagent certaines théorisations qui s'appuient sur des données psycholinguistiques et qui mettent en avant, au niveau explicatif, le rôle de différents types de mémoire (Silva et Clahsen 2008 ; Clahsen, Felsen, Neubauer, Sato et Silva 2010, *contra* Voga, Anastassiadis-Syméonidis & Giraudo 2014 ; Dal Maso & Giraudo 2014).

Présentation synthétique d'effets d'ordre morphologique issus de protocoles d'amorçage masqué grec-français

La présentation qui suit vise à donner une vision et une interprétation globales concernant les effets d'amorçage masqué inter-langue grec-français en mettant en évidence d'une part certains patterns d'effets récurrents et d'autre part des effets issus

de protocoles difficiles à mettre en place dans d'autres paires de langues. Une telle entreprise, qui ne prétend pas à l'exhaustivité mais à une prise en considération plus globale des effets d'amorçage, se fonde sur la manipulation de facteurs (variables) qui mettent l'accent sur une multitude de types de rapport entre amorces et cibles.

Les données qui sont présentées dans la table récapitulative (Table 2) ont été récoltées *via* des protocoles du même type que celui présenté plus haut, *i. e.*, amorçage masqué inter-langue dans la direction L1 vers L2, où la L1 des participants est le grec et la L2 le français. La Table 2 présente un récapitulatif des effets cognat et non-cognat trouvés dans ces protocoles, auprès de participants du même profil¹. Pour des raisons de simplicité, seule la direction d'amorçage L1 vers L2 est présentée dans le tableau. La direction opposée (L2 vers L1) a été également étudiée dans certains des travaux cités dans la table 2, mais cette direction d'amorçage, dans laquelle les effets sont d'amplitude réduite et pas systématiquement trouvés, ne sera pas abordée ici. La question des asymétries d'amorçage demeure influencée par de nombreux facteurs, entre autres des facteurs liés aux correspondances graphèmes-phonèmes (cf. Voga, 2020a: 25-26) et à la « profondeur orthographique » (*orthographic depth hypothesis*²) des langues prises en considération. De ce point de vue, une combinaison de langues comme le grec-français est à distinguer du chinois-anglais, ex., Wang et Li (2025).

Les résultats rapportés dans la table récapitulative appellent quelques précisions et remarques préliminaires :

- i. Les expériences présentées dans la Table 2 examinent des mots (ou unités lexicales, UL) cognats et non-cognats, ces deux catégories étant toujours examinées au sein de la même expérience (par ex. lignes c1, c2, c3 et d du tableau, données issues de Voga, 2020a) ;
- ii. Dans les expériences rapportées dans la Table 2, les amorces et les cibles peuvent être de différente nature : des UL morphologiquement simples (lignes a, b et b1) ou morphologiquement complexes (lignes c1, c2, c3, d, e, e1) ; des mots qui peuvent varier au niveau de leur fréquence lexicale, par ex. *idée* (ligne b) est un mot très fréquent, alors que *pluralisme* (ligne e1) l'est beaucoup moins ; des UL formées de 16 lettres, ex. αποκλειστικότητα 'exclusivité' (ligne c2) et d'autres formées par seulement 4 lettres, ex. ιδέα 'idée' ; des UL d'origine grecque (ex. *idée*) ou latine-française, ex. *rôle* (ligne b1). L'objectif ici n'est pas d'être exhaustif : il ne s'agit pas de commenter toutes les situations possibles entre grec et français, mais de montrer que

1. Sans toutefois que les participants soient les mêmes pour tous les protocoles.

2. Cette hypothèse a été proposée, à l'origine dans le cadre du fonctionnement monolingue, afin de rendre compte de différents degrés d'activation phonologique pendant le traitement, degrés qui seraient dépendants de la transparence (ou opacité) des correspondances graphèmes-phonèmes dans un système d'écriture donné. De ce point de vue, une langue comme le grec a une « profondeur orthographique » moindre par rapport à une langue comme le chinois. Cette ligne de recherche a récemment fait son entrée dans le domaine bilingue, par ex. Wang et Li (2025).

- les effets d’amorçage inter-langue sont récurrents et se manifestent au sein de toute une variété d’UL, exhibant des caractéristiques très différentes ;
- iii. Les matériaux linguistiques (stimuli) utilisés dans Voga & Anastassiadis-Symeonidis (2018, ligne e1 de la Table 2) sont issus d’une catégorie de mots qui ont la particularité d’avoir un statut intermédiaire entre l’unité construite et non construite (Corbin 1987, 457-459, « ils ne sont les produits d’aucune Règle de Construction de Mot »). Le mot d’emprunt *πλουραλιστής* /*plourali'stis*/ « pluraliste » par ex. est un mot parfaitement intégré dans le lexique grec mais dont la « base » ne signifie absolument rien et n’a pas non plus le statut d’un thème lié. Ces mots, en général pas très fréquents, peuvent avoir une petite famille morphologique, par ex. dans la famille morphologique de *ρεαλιστής* /*reali'stis*/ « réaliste », on trouve *ρεαλισμός* /*realismós*/ « réalisme » et l’adjectif *ρεαλιστικός* /*realistikós*/ « réalistique », cependant leurs « bases », par ex. *πλουραλ-* /*ploural-*/ ou *ρεαλ-* /*real-*/, ne signifient rien et ne peuvent pas être considérées comme des bases non-autonomes, ex. *°lud(e)* (Corbin 1987, 466). Les mots de ce type peuvent être considérés comme des mots complexes mais non construits, comme par ex. *royaume* ou *carpette* (Corbin 1987, 456, 458) même s’ils ont une structure interne de type [carp(ette)af]N.), *i. e.* *carp* n’appartient pas à une catégorie grammaticale et pour cette raison n’est pas une entrée lexicale.

Tableau 2 — Récapitulatif des effets d’amorçage traduction (T) et morphologique (M) obtenus dans des protocoles grec-français (direction L1 vers L2)

Type de relation entre L1 et L2	Exemples de stimuli d’amorces (gr) et cibles (fr')	Effets d’amorçage nets
a) Cognats : Forme et sens identiques (Voga 2014)	T : <i>ταξί</i> / <i>taxí</i> / ‘taxi’ - [<i>taxi</i>] M : <i>ταξιτζής</i> / <i>taxitzis</i> / - ‘chauffeur de taxi’ - [<i>taxi</i>]	T : 60 ms* M : 55 ms*
b) Cognats, forme similaire et sens identique, origine grecque b1) Cognats, forme similaire, sens identique, origine latine-française (Voga 2014 ; Voga 2020b)	T : <i>ιδέα</i> / <i>idéa</i> / ‘idée’ - [<i>idée</i>] M : <i>ιδεατό</i> / <i>idéató</i> / ‘idéa’ - [<i>idée</i>] T : <i>ρόλος</i> /‘ <i>rolos</i> / ‘rôle’ - [<i>rôle</i>] M : <i>ρολάκι</i> / <i>ro'laki</i> / ‘petit rôle’ - [<i>rôle</i>]	T : 56 ms* M : 50 ms* T : 34 ms* M : 28 ms*
c) Sens uniquement : Non-cognats c1) NC suffixés-lexicalisés (Voga 2020a) c2) NC préfixés avec base non-mot (Voga 2020a) c3) NC préfixés avec base mot (Voga 2020a)	T : <i>λογαριασμός</i> / <i>logarias'mos</i> / ‘calcul, facture’ - [<i>facture</i>] M : <i>λογαριάζω</i> / <i>loga'riazo</i> / ‘calculer’ - [<i>facture</i>] T : <i>αποκλειστικός</i> / <i>apoklistikós</i> / ‘exclusif’ - [<i>exclusif</i>] M : <i>αποκλειστικότητα</i> ‘exclusivité’ / <i>apoklistikótita</i> / ‘exclusivité’ - [<i>exclusif</i>] T : <i>σύνθεση</i> / <i>sinthesi</i> / ‘composition’ - [<i>composition</i>] M : <i>συνθέτης</i> / <i>sin'thetis</i> / ‘compositeur’ - [<i>composition</i>]	T : 41 ms* M : 21 ms NS T : 29 ms* M : 34 ms* T : -7 ms NS M : 15 ms NS

i. Les cibles sont en italiques et entre crochets. Pour les amorces, est d’abord présentée la forme écrite en grec, la transcription en forme phonique et la traduction en français.

d) Forme et sens identiques - cognats préfixés (Voga 2020a)	T: κατακλυσμός /katakli'smos/ - 'cataclysm' - [cataclysm] M: κατακλυσμαίος /kataklismi'aïos/ 'cataclysmique' - [cataclysm]	T: 57 ms* M: 50 ms*
e) Forme et sens identiques - cognats suffixés de base grecque (Voga & Anastassiadis -Syméonidis 2018)	T: μοναρχικός /monarhi'kos/ 'monarchique' - [monarchique] M: μοναρχισμός /monarhi'smos/ 'monarchisme' - [monarchique]	T: 83 ms* M: 49 ms*
e1) Forme et sens identiques - cognats suffixés 0-base (Voga & Anastassiadis-Syméonidis 2018)	T: πλουραλιστής /plourali'stis/ 'pluraliste' - [pluraliste] M: πλουραλισμός /plourali'smos/ 'pluralisme' - [pluraliste]	T: 20 ms NS M: 85 ms*

Discussion. La co-occurrence de l'effet cognat et morphologique : vers une explication morphologique ?

L'étude de la table 2 montre de manière claire la co-occurrence entre ces deux effets : il n'y a qu'un seul cas (ligne ei, cas qui sera commenté plus bas dans ce paragraphe), où l'amorçage « traduction » et « cognat » ne soient pas concomitants. Cette co-occurrence entre l'effet facilitateur induit par l'amorce « traduction » et par l'amorce « morphologique » est trouvée dans plusieurs études intra-alphabet, et notamment dans l'étude en espagnol-anglais de Sánchez-Casas & García-Albea (2005), qui fondent sur cette co-occurrence leur approche selon laquelle la relation cognat est une sorte de relation morphologique. Cette approche s'applique prioritairement, selon ces auteurs, aux langues de même alphabet (*same-script bilinguals*) comme l'espagnol-anglais. Elle est fondée sur l'idée que les cognats, en vertu de leur similarité à la fois orthographique et phonologique et de leurs caractéristiques conceptuelles, sont représentés de la même manière que les mots morphologiquement reliés au sein du lexique, alors que les non-cognats ne le sont pas, affirmation qui est mise en difficulté par nos données, comme nous le verrons plus loin. Selon ces auteurs, l'effet cognat et plus généralement la relation cognat ne doivent donc pas être vus comme le résultat d'une simple addition de la facilitation phonologique et conceptuelle, position à laquelle nous souscrivons complètement. La relation cognat devrait être vue comme un effet qui émerge de « l'architecture lexico-sémantique commune » qu'est le lexique bilingue, selon l'expression de Schoonbaert, Duyck, Brysbaert & Hartsuiker (2009), comme nous l'avons souligné dans plusieurs publications. Cette idée d'architecture commune est parfaitement compatible avec les théorisations qui, fondées sur un stockage en mémoire riche, ainsi que sur la notion d'usage (Bybee 2010, 14-32) considèrent les transferts d'activation non pas comme un mécanisme statique et unidirectionnel, mais comme quelque chose de dynamique, qui se trouve sous l'influence (entre autres) des « perceptions, conceptualisations, associations mentales » des locuteurs

(Jarvis et Pavlenko 2008, 13) et qui prend en compte l'ensemble des influences inter-langues, y compris de la L2 vers la L1.

Dans les données que nous présentons ici (Table 2) cette co-occurrence est trouvée dans tous les protocoles qui traitent de cognats, à l'exception d'un seul cas (ligne e1), celui des mots d'emprunt de type *πλουραλιστής* /plourali'stis/ « pluraliste », une catégorie de stimuli formée d'unités lexicales dont la « base » ne signifie absolument rien (appelées « mots o-base » dans l'expérience de Voga & Anastassiadis-Syméonidis 2018). Il est par conséquent assez normal que ces mots ne parviennent pas à induire un effet significatif d'amorçage « traduction », étant donné que le mot de la L1 grec et le mot français (L2) ne partagent pas d'entrée lexicale, dont la représentation serait commune, au niveau de la forme et au niveau lexical, dans un modèle de type BIA+ (Dijkstra & Van Heuven 2002 ; Peeters, Dijkstra & Grainger 2013), ou bien, aux niveaux lexical et conceptuel, dans un modèle de type RHM (Kroll *et al.* 2010). Ce résultat confirme notre hypothèse, fondée sur Corbin (1987, p. 456) selon laquelle des segments comme *πλουραλ-* /plural-/ ne constituent pas d'entrée lexicale et par conséquent ne parviennent pas à induire des effets de traduction inter-langue (Voga & Anastassiadis-Syméonidis 2018, 175), ce qui signifie entre autres, que le *locus* des effets de traduction (effets cognats) au sein du lexique mental bilingue est une représentation (commune au niveau lexical et sémantique-conceptuel) davantage attachée au lexique de la L1, en tout cas pour les mots peu fréquents que nous avons employés dans cette expérience.

Faisons néanmoins remarquer que pour cette catégorie des mots, les mots o-base, l'effet morphologique est robuste et significatif, *i. e.*, l'amorce *πλουραλισμός* /plourali'smos/ « pluralisme » induit 85 ms de facilitation sur la reconnaissance de la cible *pluraliste*. Cet effet nous oriente également vers une explication de type morphologique de l'effet cognat, puisque l'on peut l'attribuer à la connexion, forcément de type morphologique, entre le suffixe *-iste* et *-isme*, à la différence près qu'ici, dans le cadre du protocole bilingue que nous avons mis en place, cette connexion dépasse les langues (« transcends languages », comme l'écrit Bybee) de manière à ce que l'amorce *πλουραλισμός* /plourali'smos/ « pluralisme » parvienne à induire une facilitation sur la reconnaissance de la cible *pluraliste*.

À l'exception de cette catégorie de mots o-base, nous pouvons observer, dans les données rapportées sur la Table 2, qu'à chaque fois qu'un effet cognat (« traduction ») est trouvé, un effet morphologique l'accompagne. De surcroît, les conditions « traduction » ne diffèrent pas statistiquement des conditions « morphologiques¹ », ce qui nous conduit à admettre l'hypothèse de Sánchez-Casas & García-Albea (2005) et à la valider pour les données inter-alphabet grec-français. Il n'y a en effet aucune raison d'exclure de « l'explication morphologique de l'effet cognat » les cognats grec-français, qui partagent une similarité phonologique plus ou moins

1. À l'exception de la case e1.

élevée selon le cas, leurs caractéristiques conceptuelles (même signification), mais également une structure interne du même type, ex. base + suffixe dans le cas de la paire de mots ιδεατό /idéa'to/ - « idéal », ou encore μοναρχικός /monarhi'kos/ « monarchique ». Ceci est le cas également de mots comme κατακλυσμιαίος /kataklismi'aïos/ « cataclysmique » (ligne d) dont la longueur (14 lettres) pourrait empêcher la perception de la part du sujet dans le temps extrêmement bref de présentation de l'amorce. Nous constatons avec nos résultats que ceci n'est pas le cas, puisque les effets traduction et morphologique sont robustes et concomitants (57 ms* et 50 ms*, respectivement) pour ce type de stimulus. Nous adhérons par conséquent à l'explication morphologique de l'effet cognat (Sánchez-Casas & García-Albea 2005), et, comme nous le verrons plus loin, nous proposons de l'étendre à certains types de non-cognats français-grec. Ce type d'interprétation serait aussi compatible avec la position théorique de Bybee (1985 ; 1988 ; 1995), selon laquelle la connexion entre les deux lexiques (ou entre les mots du lexique unifié) serait de nature morphologique.

Avant de clore ce paragraphe, faisons néanmoins remarquer que le modèle BIA+ n'inclut pas de niveau de représentation morphologique et que l'idée de représentations morphologiques communes entre cognats est rejetée (Dijkstra, Miwa, Brummelhuis, Sappelli & Baayen 2010 ; Peeters, Dijkstra & Grainger 2013). Le type d'explication que proposent Sánchez-Casas & García-Albea (2005) et auquel nos données peuvent être intégrées est plus proche du RHM (Kroll *et al.* 2010), même si ce dernier est *a priori* un modèle de production et non pas de perception.

Le rôle de l'origine étymologique des cognats

L'explication morphologique de l'effet cognat que nous venons de voir nous conduit sur l'interrogation relative à l'origine étymologique des cognats, surtout au sein de la combinaison de langues que nous examinons ici, qui comportent de très nombreux mots d'emprunt, dans les deux sens, du grec au français et du français au grec (Anastassiadis-Syméonidis 1994). Les données de l'expérience rapportée sur la Table 2 (lignes b et b1, données venant de Voga 2020b) montrent que les deux catégories de cognats ne se comportent pas exactement de la même manière : si ces deux catégories induisent un effet traduction et morphologique, tous les deux significatifs et ne différant pas statistiquement entre eux, nous observons aussi que l'amplitude est plus large pour les cognats d'origine étymologique grecque, qui induisent 56ms d'effet traduction et 50 ms d'effet morphologique, alors que les stimuli étymologiquement français-latins induisent 34 ms et 28 ms d'effet respectivement. Sans rentrer ici dans le détail, nous pouvons souligner le lien étroit entre l'origine étymologique des mots et la variable taille de la famille morphologique (MFS, *Morphological Family Size*, de Jong, Schreuder et Baayen 2000) que nous avons explorée dans d'autres travaux (ex. Voga 2017 ; Voga, Gardani, Giraudo 2020), dont les résultats ne sont pas reportés dans la Table 2.

Ce résultat corrobore l'hypothèse du RHM selon laquelle le lien entre niveau lexical et conceptuel serait plus fort pour les mots de la L1, ayant fait l'objet d'une acquisition durant l'enfance, que pour les mots de la L2, apprises après la période critique d'acquisition du langage et à partir du mot de la L1. À l'opposé, dans le cadre d'un modèle de type BIA (ou BIA+), la prédiction d'effets différents selon l'étymologie du cognat ne paraît pas possible.

Quel statut pour les non-cognats ?

Dans les cases c, c1, c2, c3 de la table récapitulative, ainsi que dans l'expérience de Voga & Grainger (2007, exp. 2) décrite plus haut en quelque détail, les effets d'amorçage non-cognat sont présentés. Dans l'article de 2007, des non-cognats morphologiquement simples (ex. λάθος /'lathos/ « erreur »- *erreur*) sont testés : ils parviennent à induire des effets de traduction significatifs, effet de 36 ms par rapport au contrôle phonologique et de 23 ms par rapport au contrôle non-relié (cf. Table de résultats, Voga & Grainger 2007, 943). Notons que la manifestation des effets non-cognat est un résultat qui peine à être intégré dans une explication morphologique de l'effet cognat, comme celle que nous avons vue plus haut. Il convient néanmoins de davantage explorer cette catégorie de mots non-cognats, par définition très hétérogène. C'est précisément ce que l'expérience de Voga (2020a) s'est donnée comme objectif. Quatre catégories de stimuli ont été testées dans les deux directions d'amorçage (L1 vers L2 et inversement ; seuls les résultats de la première seront présentés ici) et peuvent être résumées ainsi :

- a. des paires de cognats préfixés, *e.g.*, κατακλυσμός /katakli'smos/ - *cataclysm*, dont il est connu qu'ils induisent des effets de traduction inter-langue, ainsi que des effets morphologiques (ligne d, Table 2). Ils fonctionnent par conséquent dans cette expérience comme une base (*upper baseline*) à partir de laquelle seront évalués les effets des autres catégories ;
- b. des paires de mots non-cognat qui présentent une structure interne (mots construits) : les deux mots (*i.e.*, l'amorce en grec et la cible en français) portent un préfixe avec le même sens, mais pas avec la même forme, par ex., διά /di'a/ et *trans-*, ou από /a'po/ et *ex-*, comme dans l'exemple αποκλειστικός /apoklisti'kos/ « exclusif ». La « base » de ces mots préfixés est un mot non-existant, *i.e.* qui n'existe pas comme entrée lexicale, par ex. la forme κλειστικός /klisti'kos/ « clusif » n'est pas un mot, même s'il peut contribuer à activer une famille morphologique (en grec, celle du verbe κλείνω /'kline/ « fermer »). Nous appellerons cette catégorie d'UL « NC préfixé + non-mot » (ligne c2) ;
- c. des paires de mots préfixés non-cognats ayant le même type de préfixe que ceux de la catégorie précédente (approximativement du même sens, mais sans similarité formelle) et dont la base constitue un mot dans les deux langues du bilingue, par ex. l'UL σύνθεση /'sinthesi/ « composition », qui

peut être découpée en deux parties, préfixe + base-mot, en grec comme en français. Nous appellerons cette catégorie « NC préfixe + mot » (ligne c3) ;

- d. dernière catégorie, celle de non-cognats sans préfixe, mais avec approximativement le même nombre de lettres que les mots des trois premières catégories, ex. : λογαριασμός /logaria'smos/ « calcul, facture » (ligne c1).

Le pattern de résultats dans la direction L1 vers L2 montre de robustes effets (morphologiques et traduction) pour les cognats et les « NC préfixe + non-mot » mais aucune facilitation pour la catégorie « NC préfixe + mot ». Cela est interprété comme la démonstration du fait que les effets non-cognats peuvent être déclenchés par le sens commun (au niveau de l'UL), en la présence d'une structure interne complexe « préfixe + base », à condition que cette « base » ne corresponde pas à une entrée lexicale. Ces résultats indiquent également que le système de traitement repère extrêmement vite la structure interne complexe (cf. Voga & Giraudo 2017 pour une discussion sur le rôle de la structure complexe). Cela se fait malgré l'absence totale de similarité formelle, sur la partie « préfixe » comme sur la partie « base¹ », ce qui souligne un trait fort du système, consistant à créer des amas de mots complexes sur la base du sens, en l'absence de toute similarité formelle (à partir d'un niveau avancé de compétence en L2). L'activation créée *via* ces connexions inter-langue ne parvient toutefois pas à faire surface lorsque la compétition entre unités lexicales est exercée, comme le démontrent l'absence de tout effet significatif d'amorçage pour les stimuli qui ont une base-mot. De manière schématique (pour une discussion, cf. Voga 2020a : 24-25, 29) ces données démontrent que la multiplicité des représentations activées crée de la compétition entre UL (*i.e.* du « bruit lexico-sémantique ») et désoriente ainsi le système de traitement. Ces données renforcent l'idée que les représentations lexico-sémantiques sont une composante essentielle de la co-activation des langues chez le bilingue, et ceci dès les premiers stades du traitement (Brysbaert and Duyck 2010 ; Schoonbaert *et al.* 2009). La compétition entre unités lexicales du lexique mental bilingue, indépendante de la langue à laquelle appartiennent ces dernières, suggère que dans ce processus de co-activation des langues, la compétition qui émane des bases-mots est plus importante que celle qui émane des segments qui ne correspondent pas à des entrées lexicales, et n'ont donc pas de représentation lexicale.

Ces données, combinées à celles de la catégorie opposée d'amorçage (L2 vers L1) où une asymétrie d'amorçage est trouvée, suggèrent l'existence d'un lexique mental intégré, qui contient les deux langues du bilingue. Cette architecture fait face à plusieurs types de complexité, *via* différents types de connexions, y compris celles au niveau du préfixe, suggérant que des « unités-préfixe » contribuent à grouper les mots complexes à travers les langues. La discussion souligne le rôle du sens dans la co-activation des langues et démontre que la morphologie s'applique non seulement à ce qui existe à l'intérieur du mot construit (ex. morphèmes), mais

1. Le terme o-base serait ici plus approprié.

s'étend au-delà, au niveau de la « chaîne plus large de relations morphologiques » (Mulder *et al.* 2014, 70), qui apporte ainsi un principe d'organisation du lexique bilingue.

Conclusion

L'ensemble de résultats que nous venons de présenter sont compatibles avec l'explication morphologique de l'effet cognat (Sánchez-Casas & García-Albea 2005), à condition qu'on puisse étendre cette explication à certains types de non-cognats et plus précisément à des non-cognats construits, formés d'un préfixe de sens similaire entre les deux langues et d'une base qui ne correspond pas à une entrée lexicale. Il est utile ici de relever la similitude du pattern d'effets morphologiques que nous observons entre, d'une part, les non-cognats construits de type préfixe + o-base (*i. e.*, base qui ne correspond pas à une entrée lexicale, 34 ms d'effet morphologique) et les cognats suffixés de type *πλουραλισμός* /plourali'smos/ « pluralisme »- *pluraliste* (ligne e1, 85 ms de facilitation sur la reconnaissance de la cible dans la condition d'amorçage morphologique). Cela signifierait que la présence d'une base-mot (*word-base*) n'est pas une condition *sine qua non* pour la manifestation d'un effet de type morphologique, vu que l'absence de base peut être contrecarrée par la présence de la « chaîne plus large de relations morphologiques » (Mulder *et al.* 2014: 70).

Nos résultats soulignent plusieurs faits qui sont à prendre en considération dans l'étude du traitement bilingue inter-alphabet et devraient contraindre les modélisations du lexique mental. Nous pouvons les résumer comme suit :

- i. le fait que « l'architecture lexico-sémantique commune » qu'est le lexique mental bilingue correspond à une structure intégrée, c'est-à-dire qui contient les deux langues du bilingue (« lexique unifié »), conformément à l'approche qu'est devenue dominante aujourd'hui, formulée autour de l'idée de la co-activation des langues.
- ii. cette architecture lexico-sémantique commune est construite sur la base de relations de type morphologique et morphosémantique (Corbin 1987/1991), puisque les nombreux effets non-cognats documentés ici ne font que démontrer la force de la composante sémantique durant le traitement bilingue. Une telle approche correspond à celle de Bybee (*Network Model*, Bybee 1985 ; 1995), présentée à l'origine pour le système monolingue, selon laquelle les mots sont organisés en paradigmes lexicaux structurés morphologiquement, à travers les cooccurrences entre forme et contenu. En vertu de cette organisation en paradigmes lexicaux, la présentation du cognat *ιδέα* /idéa/ « idée » ou *ρόλος* /'rolos/ « rôle » (lignes b et b1) en grec amorce la reconnaissance du cognat *idée* ou *rôle* en français (L2). Par ailleurs, la co-activation des langues implique d'une part l'impos-

sibilité de « désactiver » sa L1 (voire sa L2, sujet qui n'a pas été traité ici) et d'autre part l'activation des parents morphologiques à travers les langues du bilingue (*cross-language activation*, ex. Mulder, Dijkstra & Baayen 2015). Bybee présente le réseau de relations à la fois formelles et sémantiques qui constitue la base du lexique mental selon les termes suivants :

Les mots qui entrent dans le lexique sont reliés à d'autres mots à travers des ensembles de connexions entre des caractéristiques phonologiques et sémantiques identiques. Ces connexions entre les items ont comme effet d'induire une analyse morphologique interne des mots complexes [...] Même si les mots qui entrent dans le lexique ne sont pas scindés en leurs morphèmes constituants, leur structure morphologique émerge des connexions qu'ils établissent avec d'autres mots dans le lexique. Des ensembles parallèles de connexions phonologiques et sémantiques, s'ils sont répétés à travers des multiples ensembles de mots, constituent les relations morphologiques. (Bybee 1995 : 428).

Il est à noter que pour Bybee ainsi que pour d'autres représentants de la Morphologie Constructionnelle (*Constructional Morphology*, Corbin 1987/1991 ; Booij 2010 ; 2015), la conscience et analyse morphologiques ne correspondent pas du tout à la logique du découpage en morphèmes. Pour reprendre les mots de Bybee (1988, 127) : « Quand un nouveau mot morphologiquement complexe est appris, il forme des connexions avec le matériel existant sur la base de son sens et de son schéma phonologique. Le mot n'est pas physiquement démembré (*physically dismembered*), mais ses parts sont néanmoins identifiées. »

Il s'agit donc de la construction d'un système de correspondances qui émerge, et non pas d'une décomposition plus ou moins « obligatoire » en morphèmes, comme le soutient une approche qui a eu son heure de gloire en matière de traitement morphologique (ex. Rastle, Davis & New 2004 ; concernant le traitement bilingue, ex. Coughlin & Tremblay 2015).

- iii. les effets d'amorçage en général, et en particulier les effets d'amorçage inter-langue que nous avons étudiés ici, correspondent à la somme finale d'activation/inhibition, autrement dit, à un ensemble d'influences qui proviennent de différentes sources : la similarité formelle, le rapport sur le plan du sens, la lexicalité de la base, l'origine L1 et/ ou L2, l'origine étymologique (ou la Taille de la famille morphologique), et bien évidemment la fréquence — un facteur que nous avons « mis à plat » dans les expériences présentées ici. Ceci implique que l'étude de l'organisation du lexique mental (monolingue ou bilingue) passe par celle des connexions en son sein, qui peuvent également induire de la compétition, comme nous l'avons vu plus haut. Il convient d'admettre que cette compétition entre unités lexicales serait difficile à imaginer s'il n'y avait aucune connaissance emmagasinée dans le lexique mental. Parmi les modélisations compatibles avec des processus tels que la compétition au sein du lexique mental, nous pouvons

distinguer celle de Pirrelli, Ferro & Marzi (2015) qui adopte une approche non-modulaire sur la morphologie, afin de proposer un modèle computationnel fondé sur l'hypothèse que la connaissance des mots emmagasinée (*stored word knowledge*) et le traitement en ligne reposent sur un socle commun de principes liés à l'auto-organisation de la mémoire. Ce type de déploiement computationnel (*implementation*) est fondé sur l'idée en faveur de la quelle nous oriente l'examen synthétique de nos données : le lexique mental (bilingue comme monolingue) est un système hautement dynamique, fondé sur des principes comme l'auto-organisation en fonction de la compétence du sujet (en L1, L2, etc.), l'émergence d'ensembles de connexions dans le sens de Bybee (1995) et des processus comme la compétition, signe, entre autres, du rôle que joue la mémoire liée aux mots.

Références bibliographiques

- ANASTASSIADIS-SYMÉONIDIS, Anna, *Emprunt néologique du grec moderne. Emprunts directs du français et de l'angloaméricain. Une analyse morphophonologique / Νεολογικός δανεισμός της νεοελληνικής. Άμεσα δάνεια από τη γαλλική και αγγλοαμερικανική. Μορφοφωνολογική ανάλυση*, Thessalonique : Université Aristote, 1994.
- BRYSSAERT, Marc et WOUTER Duyck, « Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? », *Bilingualism: Language & Cognition*, 133 (2010) : p. 359-371.
- BYBEE, Joan, *Morphology: A study of the relation between meaning and form*, Amsterdam : John Benjamins, 1985.
- BYBEE, Joan, *Morphology as lexical organisation*, dans M. HAMMOND & M. NOONAN (éd.), *Theoretical Morphology. Approaches to modern linguistics*, San Diego : Academic Press, 1988, p. 119-142.
- BYBEE, Joan, « Regular morphology and the lexicon », *Language and Cognitive Processes*, 10.5 (1995) : p. 425-455.
- BYBEE, Joan, *Frequency of use and the organisation of the language*, New York : Oxford University Press, 2007.
- BYBEE, Joan, *Language, usage and cognition*, Cambridge : Cambridge University Press, 2010.
- CLAHSEN, Harald, Claudia FELSEN, Kathleen NEUBAUER, Mikako SATO et Renata SILVA, « Morphological structure in native and non-native language processing », *Language Learning* 60 (2010) : p. 21-43.

- CORBIN, Danielle, *Morphologie dérivationnelle et structuration du lexique* (2 vol.), Tübingen / Villeneuve d'Ascq, Max Niemeyer Verlag / Presses universitaires de Lille, 1987/1991.
- CHRISTOFFANINI, Paula, Kim KIRSNER et Dan MILECH, « Bilingual lexical representation: The status of Spanish-English cognates », *Quarterly Journal of Experimental Psychology*, 38A (1986): p. 367-393.
- COUGHLIN, Kaitlin et Annie TREMBLAY, « Morphological Decomposition in Native and non-Native French Speakers », *Bilingualism: Language and Cognition*, 18 (2015): p. 524-542.
- DAL MASO, Serena et Hélène GIRAUDO, « Masked morphological priming in Italian L2: Evidence from masked priming », *Linguisticae Investigationes*, 37.2 (2014): p. 322-337.
- DE GROOT, Annette et Gerald NAS, « Lexical representation of cognates and noncognates in compound bilinguals », *Journal of Memory and Language*, 30 (1991): p. 90-123.
- DIJKSTRA, Ton, Jonathan GRAINGER et Walter VAN HEUVEN, « Recognition of cognates and interlingual homographs: The neglected role of phonology », *Journal of Memory and Language*, 41 (1999): p. 496-518.
- DIJKSTRA, Ton, Kozi MIWA, Bianca BRUMMELHUIS, Maya SAPPELLI et Harald BAAYEN, « How cross-language similarity and task demands affect cognate recognition », *Journal of Memory and Language*, 62.3 (2010): p. 284-301.
- DIJKSTRA, Ton et Walter VAN HEUVEN, « The architecture of the bilingual word recognition system: From identification to decision », *Bilingualism: Language and Cognition*, 5.3 (2002): p. 175-197.
- DUFOUR, Robert et Judith KROLL, « Matching words to concepts in two languages: A test of the concept mediation model of bilingual representations », *Memory and Cognition*, 23 (1995): p. 166-180.
- ELLIS, Nick, « Selective attention and transfer phenomena in L2 acquisition. Contingency, cue competition, salience, interference, overshadowing, blocking and perceptual learning », *Applied Linguistics*, 27 (2006): p. 164-194.
- GRAINGER, Jonathan et Arthur JACOBS, « Orthographic processing in visual word recognition: A multiple read-out model », *Psychological Review* 103 (1996): 518-565.
- GOLLAN, Tamara, Kenneth FORSTER et Ram FROST, « Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals », *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23 (1997): p. 1122-1139.

- GROSJEAN, François, « Studying bilinguals: Methodological and conceptual issues », *Bilingualism: Language and Cognition*, 1 (1998) : p. 131-149.
- GROSJEAN, François, « The bilingual's language modes », dans J. NICOL (éd.), *One Mind, Two Languages: Bilingual Language Processing*, Oxford : Blackwell, 2001, p. 1-22.
- IAKOVLEVA, Tatiana, Anna PIASECKI et Ton DIJKSTRA, « Are you reading what I am reading? The impact of contrasting alphabetic scripts on reading English », dans V. PIRRELLI, C. MARZI, M. FERRO (éd.), *Word Structure and Word Usage, Proceedings of the NetWordS Final Conference*, Pisa, March 30-April 1, 2015.
- JARVIS, Scott et Aneta PAVLENKO, *Crosslinguistic influence in language and cognition*, New York, Abington : Routledge, 2008.
- JIANG, Nan et Kenneth FORSTER, « Cross-language priming asymmetries in lexical decision and episodic recognition », *Journal of Memory and Language*, 44 (2001) : p. 32-51.
- KROLL, Judith, Janet VAN HELL, Natasha TOKOWICZ et David GREEN, « The Revised Hierarchical Model: A critical review and assessment », *Bilingualism: Language et Cognition*, 133 (2010) : p. 373-381.
- KROLL, Judith et Erica STEWART, « Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations », *Journal of Memory and Language*, 33 (1994) : p. 149-174.
- LAVAU, Jean-Marc et Noëlle FONT, « Représentation des mots cognats et non cognats en mémoire chez les bilingues français-espagnol », *Psychologie française*, 43.4 (1998) : p. 329-338.
- LEMHÖFER, Kristin, Ton DIJKSTRA et Marije MICHEL, « Three languages, one ECHO: Cognate effects in trilingual word recognition », *Language & Cognitive Processes*, 19 (2004) : p. 585-611.
- MCCLELLAND, Jay et David RUMELHART, « An interactive activation model of context effects in letter perception: Part I. An account of basic findings », *Psychological Review*, 88 (1981) : p. 375-407.
- MCCLELLAND, Jay et Karalyn PATTERSON, « Rules or connections in past tense inflections: what does the evidence rule out? », *Trends in Cognitive Sciences*, 6.11 (2002) : p. 465-472.
- MIWA, Koji, Ton DIJKSTRA, Patrick BOLGER et Harald BAAYEN, « Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals », *Bilingualism: Language et Cognition*, 17.3 (2014) : p. 445-463.

- NAKAYAMA, Mariko, Christopher SEARS, Yasushi HINO et Stephen LUPKER, « Cross-script phonological priming for Japanese English bilinguals: Evidence for integrated phonological representations », *Language and Cognitive Processes*, 27.10 (2012) : p. 1563-1583.
- NAKAYAMA, Mariko, Rinus VERDONSCHOT, Christopher SEARS et Stephen LUPKER, « The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: Further support for the phonological account », *Journal of Cognitive Psychology*, 26. 7 (2014) : p. 714-724.
- PLAUT, David, Jay McCLELLAND, Marc SEIDENBERG et Karalyn PATTERSON, « Understanding normal and impaired word reading: Computational principles in quasi-regular domains », *Psychological Review*, 103 (1996) :p. 56-115.
- PEETERS, David, Ton DIKSTRA et Jonathan GRAINGER, « The representation and processing of identical cognates by late bilinguals: RT and ERP effects », *Journal of Memory and Language*, 68 (2013) : p. 315-332.
- PINKER, Stephen et Michael ULLMAN, « The past and future of the past tense », *Trends in Cognitive Sciences*, 6.11 (2002) : p. 456-464.
- PIRRELLI Vito, Marcello FERRO et Claudia MARZI, « Computational complexity of abstractive morphology », dans M. BAERMAN, D. BROWN & G. CORBETT (éd.), *Understanding and Measuring Morphological Complexity*, Oxford: Oxford University Press, 2015, p. 141-166.
- RATTLE Kathy, Matthew DAVIS et Boris NEW, « The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition », *Psychonomic Bulletin and Review*, 11 (2004) : p. 1090-1098.
- RUMELHART, David et Jay McCLELLAND (éd.), *On learning the past tenses of English verbs, Parallel distributed Processing, Explorations in the Microstructures of Cognition (vol. 2): Psychological and Biological Models*, MIT Press, 1986.
- SANCHEZ-CASAS, Rosa et José GARCIA-ALBEA, « The representation of cognate and non-cognate words on bilingual memory: Can cognate status be characterized as a special kind of morphological relation? », J. KROLL et A. De GROOT (éd.), *Handbook of Bilingualism: Psycholinguistic Approaches*, 2005, New York: Oxford University Press, p. 226-250.
- SCHOONBAERT, Sofie, Wouter DUYCK, Marc BRYSSBAERT et Robert HARTSUIKER, « Semantic and translation priming from a first language to a second and back: Making sense of the findings », *Memory and Cognition*, 37.5 (2009) : p. 569-586.
- SEIDENBERG, Marc et Jay McCLELLAND, « A distributed, developmental model of word recognition and naming », *Psychological Review*, 96 (1989) : p. 447-452.

- SILVA, Renata et Harald CLAHSSEN, « Morphologically complex words in L1 and L2 processing: evidence from priming experiments in English », *Bilingualism: Language and Cognition*, 11 (2008) : p. 245-260.
- VOGA, Madeleine et Anna ANASTASSIADIS-SYMÉONIDIS, « Connecting lexica in bilingual cross-script morphological processing: base and series effects in language co-activation », *Lexique*, 28 (2018) : p. 160-184.
- VOGA, Madeleine et Jonathan GRAINGER, « Cognate Status and Cross-script Translation Priming », *Memory and Cognition*, 35 (2007) : p. 938-952.
- VOGA, Madeleine, Anna ANASTASSIADIS-SYMÉONIDIS et Hélène GIRAUDO, « Does morphology play a role in L2 processing? Two masked priming experiments with Greek speakers of ESL », *Linguisticae Investigationes*, 37.2 (2014) : p. 338-352.
- VOGA, Madeleine, Francesco GARDANI et Hélène GIRAUDO, « Multilingualism and the Mental Lexicon. Insights from language processing, diachrony, and language contact », V. PIRELLI, I. PLAG et W. DRESSLER (éd.), *Word knowledge and word usage : A Cross-Disciplinary Guide to the Mental Lexicon*, 2020, Berlin: Mouton De Gruyter, p. 506-552.
- VOGA, Madeleine, « Statut de “cognat” et amorçage masqué: le rôle de l'indice orthographique et de l'étymologie à travers deux expériences grec-français / Διαγλωσσική ομοτυπία και μεταφραστική προτεραιοποίηση: ο ρόλος του ορθογραφικού δείκτη και της ετυμολογίας μέσα από δύο ελληνογαλλικά πειράματα », *Studies in Greek Linguistics / Μελέτες*, 34 (2014) : p. 104-125.
- VOGA, Madeleine, « Forme, structure morphologique et représentation dans le lexique mental bilingue: Amorçage masqué avec des cognats et des équivalents de traduction grec-français. /Μορφή, μορφολογική δομή και αναπαράσταση στο δίγλωσσο νοητικό λεξικό: Μεταφραστική προτεραιοποίηση με ελληνογαλλικά μεταφραστικά αντίστοιχα », *Studies in Greek Linguistics / Μελέτες*, 37 (2017) : p. 129-152.
- VOGA, Madeleine, « Lexical co-activation with prefixed cognates and non-cognates: evidence from cross-script masked priming », dans Marcel SCHLECHTWEG (éd.), *The learnability of complex constructions from a cross-linguistic perspective*. Series: Trends in Linguistics. Studies and Monographs [TiLSM], 345. Berlin: Mouton De Gruyter, 2020a, p. 7-38.
- VOGA, Madeleine, « Traitement bilingue inter-alphabet et morphologie: le rôle des variables paradigmatiques dans l'amorçage masqué grec-français », *Actes du 7^e Congrès Mondial de Linguistique Française CMLF 2020*, SHS Web of Conferences 78, 08001, 2020b.

- WANG, Xin, Juang WANG et Jeffrey MALINS, « Do you hear “feather” when listening to “rain”? Lexical tone activation during unconscious translation: Evidence from mandarin-English bilinguals », *Cognition*, 169 (2017), p. 15-24.
- WANG, Xin et Junmin LI, « Asymmetrical cognate facilitation effects: the orthographic depth hypothesis revisited in bi-script readers », *Bilingualism: Language and Cognition*, June 2025, p. 1-11 : <https://doi.org/10.1017/S1366728925100217>.

Les apprenants de chinois langue étrangère repèrent-ils les mots inconnus dans leur lecture ?

Une étude d'oculométrie
auprès d'apprenants francophones

Enhao LEGER-ZHENG

*Université de Toulouse II Jean-Jaurès,
Laboratoire NeuroPsychoLinguistique LNPL (UR4156)*

Lire pour apprendre

Le vocabulaire est essentiel dans une langue. La connaissance d'un mot renvoie à la maîtrise des informations lexicales à diverses dimensions (Nation, 2019). Cela implique que l'acquisition d'un mot nouveau se produit de façon incrémentielle et que le développement lexical est un processus complexe et chronophage. Dans le domaine de l'acquisition du vocabulaire, la lecture est une activité qui représente un grand intérêt. Au siècle dernier, des enseignants ont constaté que les enfants anglophones parviennent à apprendre des mots anglais nouveaux hors du parcours scolaire. Les études de Nagy et ses collègues (1985, 1987) auprès d'anglophones natifs ont fourni de solides épreuves avec des résultats prouvant que les jeunes écoliers sont capables d'inférer, grâce aux contextes, le sens d'un mot anglais non familier présent dans leur lecture puis le retenir ; ainsi, la lecture pour le loisir est, selon les chercheurs, le moyen principal pour les jeunes écoliers de développer leur vocabulaire. En langue étrangère (L2), les résultats des études réalisées auprès d'apprenants de différentes langues visées sont généralement en accord sur le fait qu'un apprenant d'une L2 parvient à apprendre de nouveaux mots via la lecture, que cette L2 soit alphabétique comme l'anglais (Malone, 2018 ; Sok et Han, 2020), le français (Pigada et Schmitt, 2006), l'italien (Aliyar et Peters, 2022) et l'espagnol (Pulido, 2007), ou partiellement ou entièrement logographique comme le japonais (Peterson, 2016) et le chinois (Qian, 2005 ; Zhou, 2020). Cela

implique la possibilité pour l'apprenant de travailler, non seulement pendant les cours traditionnels, mais également hors du temps de classe, sur le vocabulaire L2.

L'attention portée aux mots pendant la lecture

Ceci étant dit, l'acquisition de nouveaux mots *via* la lecture a lieu de façon incidente. Par rapport à l'apprentissage du vocabulaire sous instructions en classe traditionnelle, qui est considéré intentionnel, le résultat de comprendre puis retenir un mot non familier présent dans une lecture est fortuit (Hulstijn, 2013) : l'objectif de l'apprenant est la compréhension globale du texte ; le fait que l'apprenant tente d'inférer le sens d'un mot inconnu y figurant a aussi pour objectif de pallier à la lacune sémantique afin de rendre la lecture plus fluide.

Cela veut dire que, pendant une lecture, l'apprenant ne travaille pas sur tous les mots inconnus présents dans le texte. Par exemple, dans l'étude de Fraser (1999) auprès d'apprenants francophones d'anglais, l'absence de prise en compte d'un mot inconnu représente 24 % des stratégies utilisées par les participants face à ces mots. Dans l'étude de Paribakht et Wesche (1999), plus de la moitié des mots que les participants jugent inconnus ont été ignorés pendant l'accomplissement des tâches. En chinois L2, l'apprenant japonais sollicité dans l'étude longitudinale de Qian (2005) devait indiquer, à l'issue de chaque lecture, les mots qu'il jugeait inconnus présents dans les textes puis inférer le sens de ces mots. Le participant a effectué au total l'inférence du sens de 38 mots chinois inconnus, dont 10 n'ont pas été repérés par lui-même. Le participant ne réalisait alors qu'il ne connaissait pas ces 10 mots que lorsque l'expérimentateur l'interrogeait à leur sujet pendant les entretiens ultérieurs. Huang (2008) constate chez ses participants indonésiens que seuls les mots inconnus jugés importants pour la compréhension du texte attirent leur attention. Ellis (1994) indique que la condition primordiale à satisfaire dans l'apprentissage incident d'un mot est que l'apprenant ait conscience de l'existence de ce mot inconnu dans le texte qu'il lit, puis qu'il porte suffisamment d'attention à ce mot. Dans le cas opposé, un mot n'aura donc aucune chance d'être appris.

L'attention comprend, selon les travaux de psycholinguistes comme Schmidt (2001), trois niveaux : percevoir, constater et réfléchir pour comprendre. Le dernier niveau comporte les actes tels que les analyses d'informations et la compréhension de l'importance d'un élément constaté, donc un acte volontaire de la part de l'apprenant, fortement liée à l'objectif à atteindre ou à la motivation. Schmidt souligne que l'attention d'un apprenant est limitée, donc allouée de façon sélective et stratégique aux stimuli et soumise à un contrôle volontaire. Cela peut expliquer pourquoi, dans les études mentionnées ci-dessus, les apprenants ont délibérément contourné certains mots inconnus. Peu de gain, voire aucun, ne peut résulter d'un apprentissage sans attention, parce que les stimuli auxquels on ne fait pas attention ou peu attention ne se préservent au mieux que quelques secondes dans

la mémoire à court terme. L'attention est donc la condition nécessaire et suffisante pour que s'opère l'enregistrement des informations dans la mémoire à long terme (Schmidt 1995, 2001 ; VanPatten 1994).

Observation de l'attention portée par l'apprenant aux mots à l'aide de l'oculométrie

Traditionnellement, au moyen de rapports verbaux, introspectifs ou, plus souvent, rétrospectifs, les chercheurs observent si les apprenants prennent conscience de la présence de tous les mots inconnus contenus dans les textes testés, puis analysent leur comportement face à un mot inconnu. Dans la pratique, les expérimentateurs demandent aux participants, après la lecture, de marquer les mots qu'ils jugent inconnus et d'indiquer comment ils ont traité ces mots pendant la lecture. Or, des interrogations ont été posées sur la validité de cette méthode verbale, notamment dans le cas rétrospectif : les rapports verbaux reflètent-ils objectivement et intégralement ce qui se passait pendant la lecture ? Plus précisément, est-il possible que l'apprenant oublie d'avoir porté son attention, même brièvement, à un mot dans le texte pendant la lecture ? (Cf. Bowles, 2010).

L'oculométrie (*Eye-tracking*) fournit une piste méthodologique innovante pour les chercheurs ayant un doute sur la fiabilité des données cueillies par l'intermédiaire de l'entretien verbal, ou/et voulant réaliser des observations au travers des données ne pouvant être rapportées verbalement par un sujet testé. Utilisée d'abord dans le domaine de la psychologie, l'oculométrie est une technique qui suit les mouvements oculaires d'une personne lors de l'exécution d'une tâche intellectuelle. Cette technique s'appuie sur la théorie de Just et Carpenter (1980) qui indique que les mouvements oculaires d'une personne peuvent refléter ses processus cognitifs. Cela devient l'argument général des chercheurs qui emploient l'oculométrie afin d'explorer des activités cognitives humaines. L'utilisation de l'oculométrie a connu une croissance considérable, depuis ces dernières décennies, dans beaucoup de domaines, y compris dans la compréhension écrite. Depuis les années 1990, le fonctionnement le plus courant des systèmes d'oculométrie consiste à enregistrer en vidéo les réflexions des pupilles et des cornées des yeux de l'individu. Un *eye-tracker* (c'est-à-dire l'appareil d'oculométrie) qui fonctionne en vidéo comporte généralement deux composants principaux, un illuminateur qui projette des rayons infrarouges, et une caméra qui capte les réflexions des pupilles et des cornées (Cf. Holmqvist *et al.* 2011).

L'oculométrie rapporte généralement les mouvements d'yeux d'un sujet en fonction des aires d'intérêt (*Area of Interest*, désormais AOI), qui renvoient chacune à une zone prédéfinie, dans un stimulus visuel. L'oculométrie peut fournir une grande variété de données mesurant les informations comportementales du sujet testé sur différents aspects. La durée totale de fixation du regard (*Dwell time*), la

durée de la première fixation du regard (*First Fixation Duration*) et le nombre de revisite du regard sont parmi les plus utilisées.

La mesure du temps total du regard de l'individu sur une AOI spécifique est sans doute la plus utilisée par les chercheurs parce qu'on la considère comme l'indicateur direct de l'attention d'une personne: le temps total peut refléter l'effort exigé pour traiter les informations ou la profondeur d'un processus cognitif (Jegerski 2014 ; Xu et Wang 2020). Dans la compréhension écrite, par exemple, les résultats obtenus en phase post-lecture dans l'étude de Godfroid et Schmidtke (2013) dévoilent une corrélation positive entre la durée du regard de leurs participants néerlandais et le score de reconnaissance des mots anglais inconnus présents dans les textes. De surcroît, les participants ont passé en moyenne 306 millisecondes de plus pour lire les mots inconnus dont ils ont déclaré plus tard avoir une reconnaissance sûre.

En ce qui concerne la première fixation, elle renvoie à la durée du premier passage des yeux du sujet dans une certaine AOI. Cette mesure reflète le temps nécessaire pour les traitements rapides, tels que la reconnaissance et l'identification. Elle permet aux chercheurs d'observer les facteurs qui peuvent intervenir dans le processus initial d'activation de l'accès lexical à un mot, par exemple la fréquence du mot, la complexité morphologique, la propriété orthographique et le degré polysémique (Holmqvist *et al.*, 2011). Par exemple, L. Huang *et al.* (2022) ont constaté que, chez les apprenants sinophones de l'anglais, la mémoire à court terme et la durée de la première fixation de regard peuvent prédire leur façon de traiter des informations lexicales des mots inconnus intégrés dans le texte.

La revisite, quant à elle, compte le nombre de fois où les yeux d'un sujet repassent dans une AOI déjà quittée. Selon Holmqvist *et al.* (2011), les zones sémantiquement informatives et intéressantes reçoivent souvent de nombreuses revisites. Dans le cadre de la compréhension écrite en L2, la revisite pourrait être un indicateur qu'un apprenant cherche des informations contextuelles pour inférer le sens d'un mot inconnu (Godfroid *et al.*, 2013). Dolgunsöz (2016) a, pour sa part, mené une étude auprès de 72 apprenants turcs au sujet de l'inférence lexicale dans la compréhension écrite en anglais L2. Les participants ont lu un texte comportant au total 24 mots cibles, sous la surveillance de l'oculométrie, et répondu immédiatement après aux questions de compréhension concernant le texte. Les résultats oculométriques révèlent que les participants relisaient plus fréquemment les mots cibles et, par ailleurs, passent plus de temps pour les relire, quel qu'était leur niveau linguistique en anglais. Selon Dolgunsöz, les revisites impliquent la réanalyse ou la vérification dans un contexte discursif, des informations sémantiques et syntaxiques d'un mot. Toutefois, une revisite plus fréquente et plus longue sur un mot n'aboutissait pas à une meilleure reconnaissance de ce mot, d'après les résultats du post-test.

En chinois L2, l'étude de Yang *et al.* (2024), sollicitant 139 apprenants issus des pays de Sud-est d'Asie, se focalise sur l'effet de la compétence linguistique de

l'apprenant. Les résultats montrent que la compétence linguistique corrèle négativement avec la durée de la fixation, mais positivement avec la distance de saccade (c'est-à-dire la distance entre deux points de fixation du regard consécutifs). Cela signifie qu'un apprenant plus compétent lit plus vite et saute plus de caractères pendant la lecture, parce qu'à chaque fixation du regard il parvient à traiter davantage d'informations lexicales des mots situant à droite. Ces résultats sont en accord avec ceux de l'étude de Fu (2008) qui révèlent que, pendant la lecture, la distance de saccade d'un Chinois natif peut être 2,5 fois plus importante que celle d'un apprenant ayant un niveau avancé en chinois.

L'étude de Cui (2023) se penche plutôt sur l'effet de l'espace entre deux mots chinois sur les comportements de lecture des apprenants. En chinois contemporain, un mot est souvent constitué de plusieurs caractères et, dans une phrase, les caractères se placent en juxtaposition, sans espace supplémentaire entre deux mots (ex. 我说中文和法语¹). En d'autres termes, le lecteur doit identifier lui-même les unités lexicales dans une phrase en lisant tous les caractères constitutifs. Cela représente une tâche supplémentaire pour les apprenants dont la L1 est alphabétique et s'appuie sur des mots explicitement séparés l'un de l'autre à l'aide d'un espace, comme le français. Cette étude a sollicité des apprenants dont la L1 est alphabétique, ayant un niveau intermédiaire en chinois, ainsi que des Chinois natifs, pour lire des textes pour compréhension sous la surveillance de l'oculométrie. Les textes sont présentés sous quatre formats de segmentation lexicale : sans espace supplémentaire (Format 1, 我说中文和法语。), avec un espace ajouté entre deux mots (Format 2, 我说 中文 和 法语。), avec un espace ajouté entre deux caractères qui ne constituent pas ensemble un mot (Format 3, 我说 中 文 和 法 语。) et, dans le quatrième cas, avec un espace ajouté entre deux mots comme dans le deuxième cas, mais où le texte est cette fois entièrement présenté en transcription *pinyin*. (Format 4, Wǒ shuō Zhōngwén hé Fǎyǔ。). Les scores de compréhension montrent que, autant pour les apprenants L2 que pour les Chinois natifs, le format de présentation n'a pas d'impact sur le degré de compréhension des textes. Pour les natifs, la condition de lecture la plus contraignante est celle dans laquelle les textes sont présentés en *pinyin*. Les yeux des natifs sautaient moins quand ils lisaient les textes comportant des espaces supplémentaires, ce qui implique une perturbation causée par ces espaces, un élément non habituel dans la lecture naturelle pour les natifs. Au contraire, les apprenants L2 ont passé significativement plus de temps pour lire les textes présentés sous Format 1, mais le moins longtemps quand les textes sont présentés en *pinyin*. ; la première fixation du regard est plus longue sur les textes présentés sous Format 1 que sous les autres formats ; quand les apprenants L2 lisaient les textes avec les mots séparés, les nombres de fixation du regard et de recul du regard (c'est-à-dire les yeux reculent de sa position actuelle pour

1. « Je parle le chinois et le français ». → 我 (je) / 说 (parler) / 中文 (la langue chinoise) / 和 (et) / 法语 (la langue française).

relire la partie précédente du texte) diminuaient significativement. Cela indique que ces espaces artificiellement ajoutés entre deux mots ont eu un effet bénéfique sur l'identification des mots pour eux.

Étant une technique innovante qui peut objectivement refléter où, comment et combien de temps un apprenant regarde les mots dans un texte, l'oculométrie est aujourd'hui largement utilisée dans les études sur la compréhension écrite dans une grande palette de langues visées. Toutefois, nous n'avons trouvé, jusqu'au moment de notre expérience, aucune étude en chinois L2 mettant en lumière l'attention qu'un apprenant porte sur des mots inconnus qu'il rencontre dans une lecture.

Questions de recherche et protocole expérimental

Pour notre part, nous aimerions observer comment un apprenant du chinois L2 distribue son attention durant une lecture. Précisément, nous nous sommes interrogées sur les questions suivantes :

- Q1. L'apprenant porte-t-il davantage d'attention à un mot inconnu, par rapport à un mot connu, présent dans le texte ?
- Q2. L'apprenant accorde-t-il la même attention à tous les mots inconnus présents dans le texte ? Si la réponse est négative, une attention plus marquée portée à un mot inconnu particulier peut-elle favoriser une meilleure compréhension de ce mot ?

Afin de répondre à ces questions, nous avons mené une étude auprès d'apprenants francophones du chinois. Nous avons opté pour la technique d'oculométrie afin d'observer l'attention de nos participants pendant leur lecture. Comme, jusqu'à présent, nous n'avons trouvé aucun article portant sur l'utilisation de l'oculométrie pour mesurer l'attention qu'un apprenant porte aux mots non familiers lors de la compréhension d'un texte en chinois L2, cette étude peut être considérée comme exploratoire.

Participants

Nous avons sollicité au total 27 apprenants francophones du chinois¹. Au moment de l'expérience, nos participants, âgés en moyenne de 22 ans, étaient des étudiants universitaires en troisième année de Licence en France. Ils ont étudié le chinois pendant cinq semestres à l'université et, selon leurs manuels pédagogiques², leur maîtrise du vocabulaire chinois était estimée entre 600 et 800 mots. Cela correspond au niveau 3 du HSK³ (*Hanyu Shuiping Kaoshi*, test d'évaluation des compétences linguistiques en chinois chez un apprenant non natif).

1. Nous avons, au départ, sollicité 33 participants et ensuite exclu les données oculométriques de quatre participants. Ces données ne sont pas exploitables à cause de la mauvaise qualité.

2. Tome 1 et tome 2, *Le chinois contemporain* 当代中文, 2007, Peking University Press.

3. <https://www.chinesetest.cn/HSK/3> (consulté le 15 mai 2025).

Ces participants n'ont pas de problème de vue. Lors de l'expérience, ils ne portaient pas de lunettes, ni de lentilles de contact, et ils n'étaient pas maquillés aux niveaux des yeux et des paupières.

Instruments expérimentaux

Dispositifs de l'oculométrie

Dans le but de mettre à l'aise nos participants pendant l'accomplissement des tâches expérimentales, nous avons opté pour le système SMI Red250, un modèle d'eye-tracker fixe sans mentonnier, afin d'imiter, dans la mesure du possible, un contexte naturel de lecture. Le SMI Red250 fonctionne à une fréquence de 250 Hz, fixée au bord inférieur d'un écran de 22 pouces à résolution 1920 x 1080 (pixel), lequel a une fréquence de rafraîchissement de 60 Hertz. Ce dispositif a été jugé comme le meilleur compromis répondant à nos exigences, à la fois sur la précision de calibration et sur le confort du participant pendant la lecture.

Mots cibles et textes testés

Nous avons sélectionné au total 20 mots cibles (Tableau 1 ; les significations des mots cibles sont fournies par le *Dictionnaire chinois-français* 汉法词典, 2001). Au moins une semaine avant l'expérience, nous avons effectué un test du vocabulaire (*Vocabulary Knowledge Scale*, ou VKS, Paribakht & Wesche, 1993), dans le but de nous assurer que nos participants ne connaissaient pas les mots cibles sélectionnés. Lors de ce test, sur une liste de 60 mots chinois comportant les 20 mots cibles, les participants devaient indiquer, pour chaque mot, s'ils le connaissaient ; pour un mot jugé connu, les participants devaient ensuite préciser s'ils le connaissaient seulement par sa forme ou également par son sens, puis donner le sens du mot dans ce dernier cas.

Étant donné que le dispositif d'oculométrie utilisé ne comporte qu'un seul écran, nous avons choisi d'afficher l'intégralité du texte testé sur cet écran afin de mieux suivre les mouvements oculaires des participants pendant la lecture, plutôt que d'afficher une partie du texte à la fois et de passer à la suivante à l'aide du clavier ou de la souris. Dans le cas d'un affichage partiel, les participants auraient pu être perturbés par le changement d'affichage lorsqu'ils cherchaient, par exemple, des informations contextuelles au début du texte pour inférer le sens d'un mot cible situé vers la fin. Un tel affichage partiel aurait ainsi rendu le suivi oculaire plus difficile et aurait par conséquent biaisé les données d'oculométrie collectées.

Tableau 1 — Mots cibles sélectionnés

Texte A		Signification	Texte B		Signification
Ao1	节俭	économe, frugal	Bo1	安适	tranquille et confortable
Ao2	窗花	papier découpé pour la décoration d'une fenêtre	Bo2	五颜六色	multicolore

Ao3	送灶	saluer avec des offrandes le départ du Génie du foyer pour le Ciel, le 23 de la 12e lune	Bo3	明丽	clair et beau
Ao4	腊月	Le 12e mois du calendrier lunaire	Bo4	彩虹	arc-en-ciel
Ao5	汤圆	boulette de riz glutineux farcie de viande (fourrée à la confiture)	Bo5	踏青	faire une excursion (se promener) à la campagne au printemps
Ao6	雀跃	sauter de joie	Bo6	风筝	cerf-volant
Ao7	还礼	faire un présent en retour, offrir un cadeau à son tour	Bo7	蝉	cigale
Ao8	在乎	prendre qch. à cœur, faire attention à	Bo8	闲聊	bavarder
Ao9	乏味	fade, insipide	Bo9	飘零	tomber en voltigeant
Ao10	盼望	espérer, désirer	Bo10	白茫茫	d'une blancheur éclatante

Ceci dit, il est important de souligner qu’une connaissance de 90 %-95 % des mots présents dans un texte est nécessaire pour qu’un lecteur comprenne ce dernier (Hirsch 2003). Cela implique qu’en nous basant sur le seuil de 95 %, afin que le texte soit théoriquement compréhensible pour nos participants, nous devons intégrer ces 20 mots cibles dans un texte comportant environ 600 à 700 caractères¹. Compte tenu de ces facteurs, dans le but d’optimiser la précision des données oculaires en espaçant suffisamment les caractères lisiblement affichés sur l’écran (cf. la section suivante), nous avons opté pour la solution de rédiger deux textes narratifs, chacun comportant 10 mots cibles (cf. la figure 2 plus bas). Le texte A, au thème des coutumes traditionnelles chinoises à la Fête du printemps, est constitué de 318 caractères, tandis que le texte B, qui décrit le paysage aux quatre saisons d’un village, comporte 321 caractères.

Format d’affichage des stimuli d’oculométrie

Afin de surveiller puis comparer l’attention des participants portée sur les mots cibles et non cibles pendant leur lecture, nous avons opté pour la solution de définir, pour chaque caractère affiché, une AOI de dimension identique. Toutes les AOI ont été numérotées, comme le montre la figure 1, ce qui permet ainsi une analyse des données pour chaque AOI et donc pour chaque caractère. (cf. figure 1).

1. Dans *Le vocabulaire commun pour l’instruction obligatoire (ébauche)* [义务教育常用词表 (草案)], 2019], qui comporte environ 17 000 items lexicaux, les mots composés de deux caractères représentent environ 67,5 % du total. Nous avons basé nos calculs sur cette proportion.



Fig. 1 — Exemple des AOI mises en place pour les caractères.

Afin de favoriser une lecture fluide et d'optimiser la précision des données oculométriques, lors de la mise en page des stimuli affichés à l'écran, nous avons suivi les règles suivantes :

- Compte tenu de la fatigue éventuelle due à la fixation des yeux sur un écran, nous avons paramétré la couleur page en « Blanc, Arrière-plan 1, plus sombre 15 % », afin de rendre l'affichage plus agréable ;
- Les textes sont affichés à la police de caractère *Kaiti*, à la taille 14 (point), avec l'interligne de 38 (point) et l'espacement (entre les caractères) de 3,4 (point) ;
- Les mots cibles ne se situent pas aux quatre bords de l'écran ;
- Tous les caractères constitutifs d'un mot cible se situent sur une même ligne ;
- Deux mots cibles ne se trouvent pas dans une même proposition ;
- Les caractères constitutifs des mots cibles se situant dans deux lignes consécutives ne se trouvent pas dans une même colonne.

Les deux textes ont été finalement affichés, avec les mots cibles surlignés en jaune¹, sur l'écran comme dans la figure 2. Nous avons défini au total 318 AOI pour le texte A et 321 AOI pour le texte B. Les AOI sont de dimension 80 x 57 (pixel). Tous les caractères se trouvent au milieu de leur AOI qui est numérotée, comme le montre la figure 1 plus bas.

Nous avons effectué des tests pilotes en sollicitant des apprenants du même niveau en chinois que nos participants. Ces apprenants nous ont confirmé qu'ils ont pu lire les textes affichés ainsi sur l'écran sans problème, et que les deux textes sont globalement compréhensibles.

Procédure

L'expérience a eu lieu sous forme d'entretien individuel. Le jour de l'expérience, le participant s'est installé devant l'eye-tracker SMI Red250, à une distance de 55 à 75 cm. Après avoir testé la calibration d'oculométrie (les angles entre l'œil

1. Lors de l'expérience, quand les participants lisaient les textes en silence pour la compréhension, les mots cibles n'étaient pas surlignés.

du participant, les points réels de calibration affichés sur l'écran, et les points de fixation du regard du participant ne doivent pas être supérieur à 0,5°), nous avons réglé l'emplacement et la position assise du participant si nécessaire.

Nous expliquions à chaque participant qu'il s'agissait d'un exercice de compréhension et que nous lui poserions ensuite quelques questions sur le texte à l'issue de sa lecture. Nous avons également informé les participants qu'ils devaient lire le texte en silence et qu'ils pouvaient y consacrer autant de temps qu'ils le souhaitaient. Par ailleurs, nous leur avons demandé, dans la mesure du possible, de ne pas bouger la tête pendant la lecture.

Texte A

我小时候，过年真的很热闹。那时候的经济不太好，大家平时都很少买肉，只有过年时才会买很多肉。新年前一个多月，大人们就开始忙着为过年而各种准备。很多人会到红纸贴一些窗花，让家里更漂亮。新年前几天，每家都会做一些好吃的年夜饭。孩子们从腊月第一天开始就每天等着新年快点儿到，因为过年时会有新衣服穿，还可以吃糖果和很多别的好吃的。当然，过年时最让孩子们喜欢的事儿还是拿红包。很多孩子的红包最后不是自己的：父母没什么钱，别人给了孩子红包，孩子交给父母，父母再用这些钱去给其他人买东西。不过，孩子们拿到红包时都高兴极了，也不在乎钱最后是不是自己的。

现在日子好了，商店里的东西也多了，想什么时候买都可以，连发红包也能用手机。但大家都说过年越来越无聊了。以前盼望过年穿新衣服、吃好吃的、拿红包，那是一种快乐，现在都没有了。

Texte B

我的家乡是一个人口不多的小城市，交通很方便，一年四季都很温暖，是一个很舒服的地方。家乡的春天很温暖，城里的树和草坪早就换上了绿色的新装，路边都开着五颜六色的小花儿。春天也是一个多雨的季节。每场春雨过后，小城都会觉得更温暖。有时下完雨后，天上还会有漂亮的彩虹。每到天气好的时候，人们会到城市以外的地方去踏青。要是风比较大的话，还可以去小城旁边的山上放风筝。夏天很热，树上的蝉从早叫到晚。到了晚上，吃完了晚饭的人们都是去家门。大人们聚在一起打牌，孩子们就在一旁玩儿。那是一天中最热闹的时候。一到秋天，天气就开始转凉，鸟儿们都飞到温暖的南方过冬。草和树叶慢慢地变黄，秋风吹过，树叶飘落，小城安静下来。等着冬天的到来。家乡的冬天常常下雪，下雪后的小城白茫茫的，城里的草地上、树上、房子上和城外的山上都是白雪，真是漂亮极了。

Fig. 2 — Deux textes testés comportant les mots cibles.

Quand le participant était prêt, une image animée d'une croix de fixation situant en plein centre de l'écran était affichée au démarrage des stimuli de texte pour la lecture. La croix sert à guider les regards du participant au départ. L'affichage de la croix ne durait que 500 millisecondes, et au-delà elle disparaît, enchaînée par le stimulus du texte A. Une fois que le participant terminait sa lecture, il en informait oralement l'examineur, sans bouger la tête, et ce dernier faisait afficher le stimulus suivant à l'aide d'un clavier connecté. Il était demandé au participant de

répondre à deux ou trois questions sur la compréhension globale du texte. Ensuite, le texte A a été affiché à nouveau à l'écran, avec les mots cibles surlignés cette fois en jaune. Il a alors été demandé au participant d'inférer le sens des mots cibles, l'un après l'autre, avec le protocole de la Pensée-à-voix-haute (désormais « PVH », c'est-à-dire le participant doit verbaliser de façon simultanée toutes ses pensées pendant l'inférence, Ericsson, 2006). La même procédure a été répétée de façon identique pour le texte B.

Afin d'identifier les indices repérés par les participants lors de l'inférence, nous avons transcrit l'ensemble des entretiens. Pour leur analyse, nous avons adopté la typologie des sources de connaissance proposée par Paribakht et Trévillat (2007) dans leur étude sur l'inférence lexicale en anglais L2. Selon cette typologie, un apprenant peut mobiliser trois types de sources de connaissance pour déduire le sens d'un mot L2 inconnu dans un texte : non linguistique, linguistique en L1 et linguistique en L2. La dernière catégorie inclut les indices présents à différents niveaux : dans le texte, dans la phrase contenant le mot inféré et au sein du mot lui-même. En raison des particularités de la langue chinoise, nous avons enrichi la catégorie des indices situés au sein du mot inféré avec quatre sous-catégories : le sens, le son, la forme graphique d'un caractère constitutif et le mode de constitution du mot.

Résultats et discussions

L'eye-tracker SMI Red250 suit, pour chaque œil, les mouvements oculaires du sujet testé. Par le biais de la fonction « Gaze replay » du logiciel SMI BeGaze, nous avons examiné les parcours des deux yeux de chaque participant pendant la lecture en silence puis, au vu de la qualité du suivi, nous avons choisi les données de l'œil gauche pour l'analyse.

Pour rappel, nous avons défini au total 639 AOI pour les deux textes. Parmi ces AOI, 42 AOI concernent les caractères constituant les mots cibles et 597 autres AOI concernent les mots non-cibles. Nous avons exclu les données des AOI concernant quelques mots cibles pour certains participants qui les avaient déclarés connus lors du pré-test du vocabulaire. Ainsi, nous avons observé au total 17189 AOI, parmi lesquelles 16065 AOI concernent les mots non-cibles et 1124 AOI pour les mots cibles. Nous avons observé, pour chaque AOI (caractère), la durée totale du regard, le nombre de revisite du regard et la durée de la première fixation du regard. Afin d'interpréter les données de façon plus explicite, nous avons transformé le nombre de revisite du regard au nombre de visites total. C'est-à-dire, par exemple, pour l'AOI X, si le nombre de revisite du regard est de 2, le nombre de visite du regard total dans l'AOI X est alors de 3. Si aucun regard ne s'est fixé sur l'AOI X, le nombre de visite est de 0.

Il est à noter qu'il existe un décalage entre la perception oculaire d'un mot et l'activation des informations lexicales concernant ce mot dans le cerveau du lecteur.

Ce décalage est d'environ 50 millisecondes (désormais ms) chez les locuteurs natifs de l'anglais (Rayner *et al.*, 2003). Dans le cas du chinois, malgré l'absence de consensus entre les études portant sur cette question, il est généralement admis que, pour un caractère combinant la forme, le son et le sens, ces trois dimensions lexicales ne s'activent pas de manière simultanée (Chen *et al.*, 2003 ; Wang, 2007). Du reste, en chinois, le traitement des informations lexicales se produit non seulement au niveau du caractère, mais aussi au niveau du mot, constitué souvent de plusieurs caractères dans le chinois contemporain. Les résultats de Wang (2007) montrent qu'une fixation du regard de 80 ms sur un mot chinois (constitué de deux caractères) suffit à accomplir la compréhension d'un texte pour un adulte chinois. Ce seuil est de 50 ms selon la conclusion de Liu *et al.* (2015). En l'absence d'études menées auprès d'apprenants de chinois L2 au moment de notre expérience, nous avons retenu un seuil de 50 ms¹.

Lecture des textes testés

Nous avons d'abord observé les AOI pour savoir si nos participants ont sauté quelques caractères lorsqu'ils lisaient les textes. Comme le montrent les études présentées plus tôt, un lecteur, non débutant dans une langue, ne lit généralement pas tous les mots présents pendant sa lecture. Certains mots sont sautés pendant une lecture. Le taux de saut peut être impacté par les éléments tels que le niveau linguistique de l'apprenant et la compréhensibilité du texte. Un lecteur expert en anglais saute au moins 30 % des mots présents dans un texte (Angele et Rayner, 2013), et un adulte chinois peut en sauter jusqu'à 50 % (Zang *et al.*, 2017). Selon nos résultats, en moyenne pour les deux textes, 21,4 % des AOI (caractères) ont reçu une fixation du regard de moins de 50 ms. Pour le texte A, le taux de saut (19,9 %) est moins important que celui du texte B (22,8 %). Il semble, pour nos participants, le texte A est plus difficile à comprendre que le texte B.

Attention portée aux mots : mots cibles vs mots non cibles

En ce qui concerne les mots cibles et non-cibles présents dans les textes, les résultats montrent que le saut du regard sur les mots cibles est bien moins important (15,5 %) que celui sur les mots non cibles (21,8 %). Les mots inconnus semblent donc avoir attiré davantage l'attention des participants que les mots connus. Respectivement, les taux du saut sont de 15,7 % pour les mots cibles intégrés dans le texte A et de 15,2 % pour ceux du texte B, une différence négligeable.

Selon nos résultats, démontrés par la figure 3, en moyenne, les participants ont

1. Cela veut dire que, lors de l'extrait de données oculométriques par le biais de la fonction « Metrics export » du logiciel SMI BeGaze, si la durée de fixation sur un caractère (donc une AOI) est inférieure à 50 ms, nous considérons que le regard du participant sur ce caractère est trop court pour lui permettre d'entreprendre le traitement d'informations concernant ce caractère. Opérationnellement, toutes les fixations dont la durée est inférieure à 50 ms sont rectifiées à 0 ms.

fixé leur regard 44 % de plus sur les caractères constituant les mots cibles (AOI MC, durée totale de 1 099 ms) que sur ceux des mots non-cibles (AOI non MC, durée totale de 761 ms). Concernant le nombre de visite du regard, les yeux des participants ont visité en moyenne 4,6 fois les AOI contenant les mots cibles, contre 3,7 fois pour les mots non cibles, soit 24 % de plus. Quant à la première fixation, la durée en moyenne est de 473 ms sur les caractères constitutifs des mots cibles, 29 % de plus que celle sur les mots non cibles (366 ms).

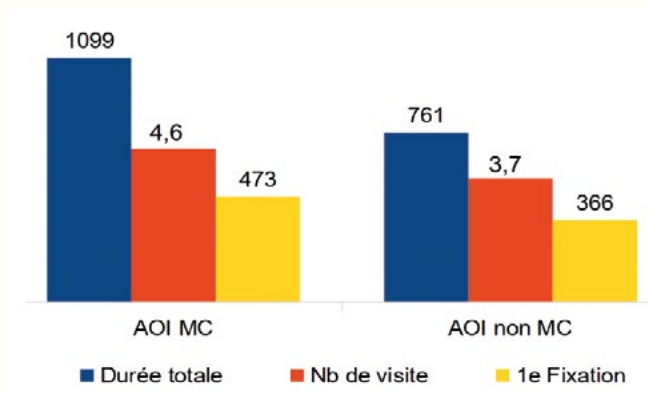


Fig. 3 — Résultats oculométriques moyens pour les AOI correspondant aux mots cibles et aux mots non cibles.

Nos résultats sont en accord avec ceux de l'étude de Godfroid et Schmidtke (2013), dans laquelle les apprenants ayant un niveau avancé en anglais L2 ont passé plus de temps à lire les mots inconnus présents dans les textes. Nos résultats révèlent que, globalement, nos participants ont bien pris conscience de la présence des mots inconnus dans les textes, et que ces mots ont exigé de leur part des traitements cognitifs plus soutenus. Si une lecture plus longue des mots inconnus constitue un résultat relativement attendu, il est intéressant de constater que les participants ont également relu plus fréquemment les caractères constitutifs des mots cibles. Rappelons qu'en chinois contemporain, l'absence d'espaces entre les mots, lesquels peuvent être formés d'un ou de plusieurs caractères, ne permet pas de marquer clairement les frontières lexicales. Pour les apprenants dont la langue maternelle (L1) est alphabétique, comme le français, l'identification des unités lexicales dans une chaîne continue de caractères représente donc une difficulté supplémentaire lors de la lecture et de la compréhension. Cette observation rejoint les conclusions de l'étude de Cui (2023), selon lesquelles les apprenants du chinois L2 présentent un nombre de fixations et de régressions oculaires significativement plus élevé lorsqu'ils lisent des textes présentés dans leur forme naturelle (sans espaces supplémentaires), comparativement à des textes où les mots sont artificiellement séparés. Ce résultat montre que la sensibilité à la composition

lexicale, autrement dit la conscience morphologique, joue un rôle déterminant dans la compréhension d'un texte en chinois. En effet, la conscience morphologique d'un apprenant est un prédicteur significatif de sa performance en compréhension écrite, que ce soit en chinois L2 (Zhang *et al.*, 2021), en chinois langue d'héritage (Zhang et Koda, 2018) ou encore en chinois L1 (Zhang, 2016). Il semble donc logique de penser que les relectures plus fréquentes des caractères constitutifs des mots cibles observées dans notre étude visent non seulement à rechercher des indices contextuels permettant de déduire le sens d'un mot inconnu (Dolgunsoz, 2016), mais aussi à identifier le mot lui-même afin de mieux comprendre la phrase dans laquelle il apparaît (Xiang *et al.*, 2018).

Les deux figures suivantes présentent en détail les résultats obtenus pour les mots cibles. Rappelons que les 20 mots cibles ne sont pas tous constitués de deux caractères. Pour cette raison, la durée totale du regard sur un mot cible a été calculée comme la moyenne des durées de fixation sur les caractères qui le composent. Par exemple, dans le cas où un mot cible M est constitué de deux caractères, si la durée totale sur son premier caractère est de 100 ms et de 200 ms sur le deuxième caractère, la durée totale du regard passée sur le mot M est alors de 150 ms. Le nombre de visite et la durée de la première fixation du regard sur chaque mot cible ont été calculés selon le même principe.

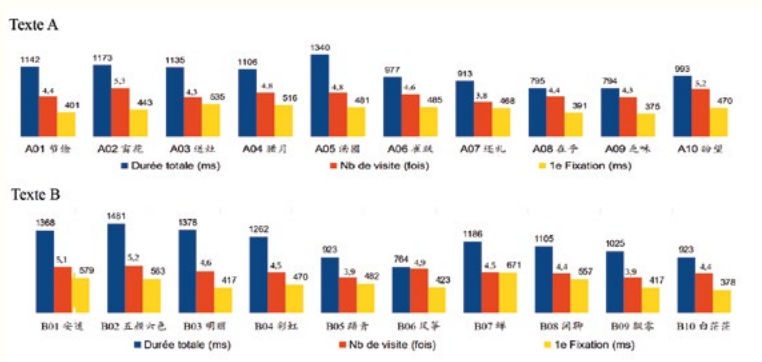


Fig. 4 — Résultats oculométriques en moyenne des mots cibles.

La figure 5 ci-dessous illustre la durée moyenne de la première fixation observée sur les caractères composant les vingt mots cibles. Nous rappelons que la première fixation du regard constitue un indicateur spécifiquement lié à la difficulté rencontrée immédiatement après la lecture d'un mot dans le texte. Pour un apprenant du chinois L2, cette difficulté est généralement associée au déchiffrement des caractères.

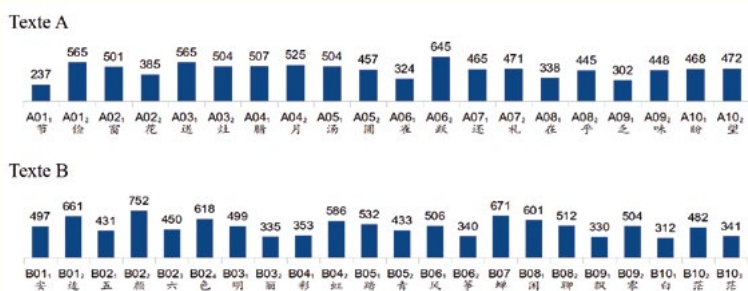


Fig. 5 — Durée de la 1^{re} fixation du regard (ms) sur les caractères constitutifs des mots cibles.

Comme le montre la figure 4, les mots cibles ont en moyenne suscité des regards plus longs et plus fréquents de la part des participants. Toutefois, certains de ces mots ne présentent pas de différences notables, en termes de mesures oculométriques, par rapport aux mots non cibles dans les textes testés. Par exemple, pour les mots cibles A08 (在乎 « prendre quelque chose à cœur »), A09 (乏味 « fade ») et B06 (风筝 « cerf-volant »), la durée totale du regard observée chez nos participants était en moyenne de 795 ms, 794 ms et 764 ms, respectivement. Par rapport à la moyenne de la durée totale du regard sur les mots non cibles présents dans les textes (761 ms), il semble que nos participants n'aient pas accordé davantage d'attention à ces trois mots inconnus par rapport aux autres mots qu'ils connaissaient. Il serait donc raisonnable d'en déduire que tous les mots cibles n'ont pas suscité la même attention de la part des participants.

Il apparaît que les éléments suivants ont influencé la manière dont nos participants ont lu ces mots.

Emplacement du mot dans le texte

Nous pouvons constater, dans la figure 4, que nos participants ont lu plus longtemps les cinq premiers mots dans chaque texte, et qu'à contrario leurs yeux ont passé moins de temps sur les cinq autres mots cibles présents dans la partie postérieure du texte. Précisément, dans le texte A, la durée totale du regard en moyenne sur les caractères constitutifs des cinq premiers mots cibles (c'est-à-dire A01, A02, A03, A04 et A05) est de 1179 ms, mais seulement de 894 ms sur ceux des cinq derniers mots (donc les mots A06~A10). Pareillement, dans le texte B, la moyenne pour les cinq premiers mots (B01~B05) est de 1314 ms, mais de 974 ms pour les autres mots cibles (B06~B10) (cf. Figure 6). Les différences entre les deux groupes de mot, en fonction de leur emplacement dans les textes A et B, sont statistiquement significatives ($p < ,001$). En revanche, au niveau du nombre de visites du regard et de la durée de sa première fixation, les différences entre ces deux groupes de mots ne sont pas significatives pour le texte A, ni pour le texte B ($p > 0,10$).

1. Test-T pour échantillons appariés.

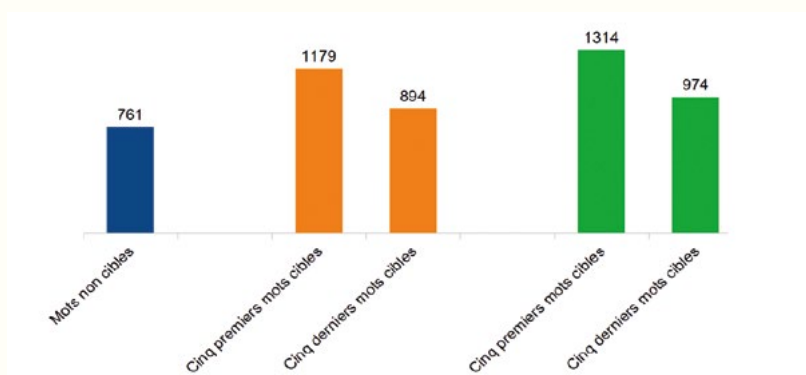


Fig. 6 — Durée totale du regard (ms) sur les mots non cibles présents dans les textes et sur les deux groupes de mots cibles intégrés dans les textes A et B.

La première fixation du regard plus longue, ainsi qu’une relecture plus fréquente des cinq derniers mots cibles par rapport aux mots non cibles, sont deux constats qui peuvent indiquer, selon nous, que nos participants ont réalisé que ces mots leur étaient inconnus lorsqu’ils les rencontraient pendant la lecture. Toutefois, leurs yeux ont passé moins de temps au total sur ces mots inconnus situés dans la partie postérieure du texte. En d’autres termes, les participants ont consacré moins de temps à ces mots lors de leur relecture. Ils semblent donc moins « patients » pour décrypter ces derniers mots inconnus. Cela est, selon nous, probablement lié à la fatigue de nos participants dans leur effort de compréhension du texte, en particulier pour les mots inconnus, ainsi qu’au fait qu’ils privilégient la compréhension globale du texte plutôt que celle des mots isolés. Il ne faut pas oublier que l’objectif de leur lecture, rappelé précédemment, est de comprendre le texte afin de répondre aux questions qui suivent. Au début de la lecture, il leur est important de collecter un maximum d’informations pour se faire une idée générale du thème. Au fur et à mesure que cette idée se précise, il devient probablement moins nécessaire pour eux de connaître le sens de tous les mots pour poursuivre la lecture. Par conséquent, ils consacrent moins d’efforts cognitifs. Le déclin de leur attention envers les mots cibles situés dans la partie postérieure du texte est peut-être l’une des raisons pour lesquelles certains caractères constitutifs, censés être inconnus des participants, n’ont pas suscité de première fixation prolongée, tels que 雀 (324 ms, dans Ao6 雀跃 *sauter de joie*), 乏 (329 ms, dans A10 乏味 *fade*), 箏 (340 ms, dans Bo6 风筝 *cerf-volant*) et 飘 (330 ms, dans Bo9 飘零 *tomber en voltigeant*) (cf. la figure 5).

Forme lexicale du mot

Parmi les 20 mots cibles, en termes de durée totale, le mot Bo2 (五颜六色 *multi-couleur*) a reçu en moyenne 1481 ms de regard, presque le double de la moyenne

des mots non cibles (761 ms) et 36 % de plus que la moyenne des mots cibles (1099 ms). L'une des raisons probables est que la forme de ce mot a intrigué nos participants : ce mot comporte deux caractères représentant des chiffres, 五 (*cinq*) et 六 (*six*), bien courants et connus par les participants ; la majorité des mots en chinois contemporain sont constitués de deux caractères, or, les deux chiffres pourraient facilement faire douter aux participants d'un lien particulier entre les quatre caractères, 五*六* (*cinq * six **). Il n'est donc pas étonnant que les quatre caractères constitutifs de Bo2 aient été fréquemment relus par nos participants (5,2 fois en moyenne), comparativement à une moyenne de 4,6 pour les mots cibles et de 3,7 pour les mots non cibles. Plus précisément, les données oculométriques révèlent que le caractère 颜 a reçu la fixation la plus longue (2 279 ms) parmi tous les caractères constitutifs des 20 mots cibles, et ce constat vaut également pour la première fixation du regard (752 ms). La première fixation sur le quatrième caractère 色 du mot (618 ms) est également longue, soit 30,7 % de plus que la moyenne des caractères constitutifs des mots cibles (473 ms). En revanche, la première fixation est de 431 ms sur 五 (*cinq*) et de 450 ms sur 六 (*six*). Cela correspond au fait que nos participants connaissent, selon leurs manuels pédagogiques, ces deux caractères de chiffre, mais non 颜 et 色. Cependant, lors de l'inférence des mots cibles avec la PVH, ils ont majoritairement reconnu le caractère 色, tandis que personne n'a reconnu 颜. Les enseignants de nos participants ont précisé qu'ils avaient abordé le thème des couleurs dans une leçon du tome 1 consacrée à l'achat de vêtements au magasin. Le mot 颜色 apparaît dans la liste des mots référents à la fin de cette leçon (mots supplémentaires, non obligatoires à maîtriser). Il est probable que nos participants aient pu retenir le caractère 色 en pratiquant les mots de couleur dans cette leçon, tels que 白色 (*blanc*) et 红色 (*rouge*), tandis que 颜, seulement abordé dans le mot 颜色, ait été peu utilisé. Dans ce cas, le fait que nos participants aient porté davantage d'attention à 颜 et 色 peut également résulter d'une autre possibilité : le caractère 色, enfin reconnu par nos participants, aurait vaguement activé dans leur mémoire des informations lexicales concernant 颜, et, par la suite, ils auraient tenté d'activer davantage d'informations sur 颜 à travers 色 ou les composants de 颜.

Un autre mot cible qui comporte plus de deux caractères est Bio (白茫茫 *d'une blancheur éclatante*). Les participants n'ont pas porté beaucoup d'attention au premier caractère 白 (*blanc*), qui leur est connu : la première fixation du regard sur ce caractère est de 312 ms, moins que la moyenne pour les caractères constituant les mots non cibles (366 ms). En revanche, ils se sont rendu compte que le caractère suivant, 茫, leur était inconnu : la première fixation a duré 482 ms, soit 54,5 % de plus que celle sur 白. Les participants n'ont pas passé autant de temps sur le troisième caractère du mot, 茫, puisqu'il est identique au deuxième. Cependant, ils ont relu plusieurs fois les trois caractères, notamment les deux premiers, respectivement 5, 4,6 et 3,5 fois. Selon nous, cela suggère qu'ils cherchaient des indices pour déterminer si ces trois caractères formaient une seule unité lexicale.

Il est intéressant pour nous de voir qu'en comparaison entre les deux mots Bo2 (五颜六色 *multicolore*) et B10 (白茫茫 *d'une blancheur éclatante*), ayant chacun une forme lexicale « atypique », nos participants ont passé moins de temps pour lire B10 que Bo2, à savoir 923 ms pour B10 contre 1 481 ms pour Bo2. Il est possible que cela s'explique par la raison mentionnée précédemment : B10 se trouvant à la fin du texte B, qui décrit le paysage des quatre saisons dans l'ordre chronologique, les participants ont probablement compris que cette partie évoque un village enneigé en hiver. De plus, ils connaissent le premier caractère 白 (blanc) de B10, qui semble cohérent avec la couleur de la neige. Les participants semblent donc s'être contentés d'une compréhension globale du texte, sans approfondir le sens du mot B10.

Composants de caractère

Concernant la durée de la première fixation, parmi les mots cibles situés dans la partie postérieure de chaque texte, le mot Bo7 (蝉, « cigale »), constitué d'un seul caractère, se distingue : il a reçu en moyenne un regard de 671 ms, l'une des durées les plus longues parmi les 42 caractères constitutifs des mots cibles, soit presque le double de la moyenne des caractères des mots non cibles dans les textes (366 ms). Pour un autre caractère, 虹 (dans Bo4 彩虹, « arc-en-ciel »), contenant également le radical 虫 (insecte), les participants ont passé en moyenne 1 749 ms à le lire, la deuxième durée la plus longue. La première fixation sur ce caractère était de 586 ms, et leurs yeux y sont revenus en moyenne 5,4 fois. Selon les entretiens avec la PVH lors de l'inférence du sens de ces mots, les participants s'inspiraient de ce radical pour essayer de déduire le sens des mots (exemple Bo, 蝉, « cigale ») : « J'aurais dit un insecte, parce que la partie gauche du caractère est la “bête”. Généralement, on la retrouve dans les caractères... comme celui du “moustique” (蚊)... »

Cela montre que nos participants sont conscients du rôle sémantique des radicaux-clés dans les caractères et qu'ils s'en servent pour deviner le sens d'un caractère inconnu. Cela est, selon les entretiens verbaux, aussi le cas de 跃 dans Ao6 (雀跃, « sauter de joie ») et de 踏 dans Bo5 (踏青, « faire une excursion à la campagne au printemps »), qui comportent tous deux le radical 足 (pied).

De plus, les durées de première fixation sur les caractères 俭 (565 ms, dans Ao1 节俭, « économe, frugal »), 腊 (507 ms, dans Ao4 腊月, « le 12^e mois du calendrier lunaire ») et 适 (661 ms, dans Bo1 安适, « tranquille et confortable ») sont relativement longues. Selon l'entretien ci-dessous, ces caractères ont rappelé aux participants d'autres caractères présentant une forme similaire (exemple Ao1, 节俭 « économe, frugal ») : « Et le deuxième (caractère) je dirais qui a la tendance à se prononcer “qian”... enfin, du moins “i-a-n” à la fin, puisque... on le retrouve dans le “lian” (脸), dans le visage et dans “examiner” (检). » Il paraît raisonnable d'induire que, chez un apprenant de niveau intermédiaire en chinois, comme nos participants, un composant peut activer plusieurs caractères le contenant et que l'apprenant connaît. De plus, l'apprenant est conscient du rôle de ce composant

dans différents caractères et est capable de l'utiliser pour extraire des informations lexicales.

Rapport entre l'attention portée sur un mot cible et le degré de la compréhension du mot

Nous avons observé au total 536 cas d'inférence. Les réponses de l'inférence sur le sens des mots cibles ont été notées selon une échelle de cinq niveaux : 0 point pour les inférences abandonnées ou complètement fausses ; 1 point pour les réponses complètement correctes ; 0,25 point, 0,5 point et 0,75 point pour les réponses partiellement correctes (en fonction de leur proximité avec la réponse entièrement correcte).

Nos résultats (figure 7) dévoilent que, le score de l'inférence en moyenne des mots cibles étant de 0,21 point, le score en moyenne des mots cibles intégrés dans le texte A est de 0,18, tandis que celui du texte B est de 0,25, soit 38,9 % de plus que celui du texte A. Cette différence démontre que nos participants ont eu plus de difficultés pour comprendre les mots cibles intégrés dans le texte A. Il semble que cela fait aussi écho à nos résultats oculométriques qui montrent que nos participants ont passé davantage de temps pour lire le texte A. Rappelons que le texte A porte sur les coutumes traditionnelles chinoises lors du Nouvel an chinois, et que le texte B est un descriptif du paysage d'un petit village aux quatre saisons. Nos participants ont l'air d'avoir eu plus de difficultés pour comprendre un mot, ainsi qu'un texte, qui exige des connaissances socio-culturelles chez l'apprenant.

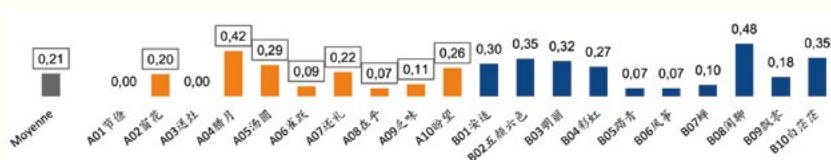


Figure 7 — Score de l'inférence du sens des mots cibles intégrés dans le texte A et dans le texte B.

Selon l'analyse des entretiens avec la PVH, lorsque nos participants tentaient de déduire le sens des mots cibles, ils s'appuyaient davantage sur les indices phrasiques que sur ceux provenant du texte global ou du mot lui-même. Cependant, les résultats statistiques montrent que, parmi toutes les sources de connaissance, seule la connaissance du sens des caractères constitutifs d'un mot cible chez les participants exerce un effet significatif et positif sur la compréhension de ce mot ($p < ,001$). En revanche, aucune des trois mesures oculométriques (durée totale du regard, nombre de visites du regard et durée de la première fixation) n'a d'impact significatif sur le degré de réussite de l'inférence du sens de ces mots ($p > 0,10$). Ces résultats indiquent que l'effort d'attention porté sur un mot chinois inconnu au

cours de la lecture n'entraîne pas systématiquement une meilleure compréhension de celui-ci. En revanche, la connaissance des caractères constitutifs par l'apprenant peut directement influencer le succès de l'inférence du sens.

Conclusions et réflexions

Nous avons observé, au moyen de la technique d'oculométrie, l'attention portée par des apprenants L2 francophones de niveau intermédiaire sur des mots chinois inconnus présents dans la lecture. À travers les mesures de la durée de la première fixation, du nombre de visites du regard et du temps total passé sur chaque mot, nos résultats révèlent la manière dont l'apprenant prend conscience de la présence d'un mot inconnu dans le texte. L'apprenant consacre davantage de temps à la lecture des mots chinois inconnus qu'à celle des mots connus, ce qui suggère que ces mots inconnus requièrent un traitement cognitif plus appuyé.

Toutefois, l'apprenant ne porte pas son attention de façon identique à tous les mots inconnus qu'il rencontre dans le texte. Un mot contenant des caractères avec des composants/radicaux familiers, ou encore à forme atypique, a tendance à attirer plus facilement son regard ; l'apprenant peut en revanche ignorer certains mots inconnus, notamment ceux présents dans la partie postérieure du texte. La raison probable est que l'objectif de la lecture est la compréhension globale du texte. Afin d'économiser ses ressources cognitives, l'apprenant ne cherche pas à comprendre tous les mots inconnus présents dans le texte lorsqu'il pense s'être fait une idée claire sur le thème de ce dernier.

L'effort d'attention que l'apprenant porte particulièrement à un mot inconnu pendant la lecture n'aboutit pas systématiquement à une meilleure compréhension de ce mot. Cette dernière est plutôt en corrélation positive avec la connaissance sur le sens des caractères constitutifs de ce mot chez l'apprenant.

À partir de nos résultats, ainsi que de ceux d'autres études montrant un effet bénéfique de la lecture sur l'apprentissage du vocabulaire, nous pensons qu'un enseignant de chinois L2 peut utiliser l'activité de compréhension écrite comme outil auxiliaire, que ce soit en classe ou à titre de devoir à la maison, pour aider les apprenants de niveau intermédiaire et plus à développer leur vocabulaire à un rythme soutenu, mais sous certaines conditions.

L'enseignant doit avant tout préparer les textes de compréhension avec soin, afin qu'ils ne soient pas surchargés de mots inconnus. Comme les apprenants s'appuient largement sur les indices contextuels pour déduire le sens d'un mot inconnu, les textes doivent rester globalement compréhensibles pour le public ciblé. Dans le but d'inciter davantage les apprenants à prêter attention aux nouveaux mots, l'enseignant doit, autant que possible, les placer dans la partie antérieure des textes.

Étant donné que les apprenants ne prêtent pas attention à tous les mots inconnus rencontrés dans un texte, il est nécessaire que l'enseignant les « incite » à le faire.

Laufer et Hulstijn (2001) soulignent que, dans le cadre de l'apprentissage incident du vocabulaire, l'un des facteurs susceptibles de prédire le gain d'un mot donné est le besoin de l'apprenant de connaître ce mot. Ce besoin relève de la dimension motivationnelle, ou de la volonté de l'apprenant de répondre aux exigences d'une tâche. Il peut provenir de l'apprenant lui-même ou être généré par un agent externe, comme les exercices. Ainsi, il est important que l'enseignant mette en valeur les nouveaux mots visés, par exemple en les liant à des questions de compréhension. L'enseignant peut, bien sûr, demander directement aux apprenants de travailler sur un mot donné présent dans un texte en le rendant concrètement saillant, par exemple en le soulignant et en leur demandant d'en déduire le sens à partir du contexte. Toutefois, l'accessibilité croissante aux outils électroniques pour les apprenants peut remettre en question l'efficacité de cet exercice. À notre connaissance, aucun outil d'intelligence artificielle ne permet actuellement de repérer automatiquement les mots inconnus pour un public ayant un niveau donné dans une langue, qu'elle soit maternelle ou étrangère. Dans ce cas, il semble préférable de demander aux apprenants de souligner eux-mêmes les mots nouveaux dans leurs textes, ce qui les oblige à effectuer une lecture plus attentive afin d'identifier les mots inconnus. Par ailleurs, selon nos résultats présentés plus haut, les mots à forme atypique, par exemple les *chengyu* (成语, proverbes ou expressions figées) composés majoritairement de quatre caractères, et ceux contenant des composants familiers dans leurs caractères constitutifs attirent davantage le regard des apprenants, facilitant ainsi le repérage des mots qu'ils ne connaissent pas.

Enfin, l'attention qu'un apprenant porte à un mot pendant la lecture n'a pas d'impact direct sur la compréhension de ce mot. L'inférence du sens à partir d'indices de toute nature ne conduit pas non plus systématiquement à une réponse correcte, comme le montrent les résultats de notre étude (figure 7). Il est donc indispensable de rendre aux apprenants le sens exact des mots visés, afin d'éviter que des informations erronées à leur sujet ne se fossilisent (Yu et Boers, 2023).



Nous adressons nos vifs remerciements à M^{me} Émilie Massa, chercheuse et enseignante en sciences du langage à l'université de Toulouse Jean-Jaurès, et M. Paul-Vincent Paubel, ingénieur d'étude CNRS, pour leurs conseils précieux pendant la conception et la mise en place des manipulations en oculométrie, ainsi que lors du traitement de données oculométriques. Nous remercions sincèrement M. Saïd Jmel, appui en statistiques aux projets scientifiques à l'université de Toulouse Jean-Jaurès, pour son aide précieuse sur les analyses statistiques.

Bibliographie

- ALIYAR, Mahnaz et Elke PETERS, « Incidental Acquisition of Italian Words from Comic Books », *Reading in a Foreign Language*, 34.2 (2022) : p. 349-377.
- ANGELE, Bernhard et Keith RAYNER, « Processing the in the Parafovea : Are Articles Skipped Automatically? », *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39.2 (2013) : p. 649-662.
- BOWLES, Melissa A., *The Think-Aloud Controversy in Second Language Research*. New York : Routledge, 2010.
- CHEN, Baoguo 陈宝国, Lixin WANG 王立新, et Danling PENG 彭聃龄, « 汉字识别中形音义激活时间进程的研究 (II) » [The time course of graphic, phonological and semantic information processing in Chinese character recognition (II)], *心理学报 [Acta Psychologica Sinica]*, 35.5 (2003) : p. 576-581.
- CUI, Yaqiong, « Eye Movements of Second Language Learners When Reading Spaced and Unspaced Chinese Texts », *Frontiers in Psychology*, 14 (2023).
- DOLGUNSÖZ, Emrah, « Using Eye-Tracking to Measure Lexical Inferences and Its Effects on Reading Rate during EFL Reading », *Journal of Language and Linguistic Studies*, 12.1 (2016) : p. 63-78.
- ELLIS, Nick, *Implicit AND Explicit Language Learning: Their Dynamic Interface and Complexity*, dans P. REBUSCHAT (éd.), *Implicit and Explicit Learning of Languages*, 1994.
- FRASER, Carol, « Lexical Processing Strategy Use and Vocabulary Learning through Reading », *Studies in Second Language Acquisition*, 21.2 (1999) : p. 225-241.
- FU, Yuping 付玉萍, « 以汉语为第二语言的留学生高级阶段阅读眼动研究 » [Investigation to the advanced Chinese reading patterns of foreign students by eye tracker], Master Thesis, 首都师范大学 [Capital Normal University], 2008.
- GODFROID, Aline et Jens SCHMIDTKE, « What Do Eye Movements Tell Us about Awareness? A Triangulation of Eye-Movement Data, Verbal Reports and Vocabulary Learning Scores », dans J. M. BERGSLEITHNER, S. N. FROTA et J. K. YOSHIOKA (éd.), *Noticing and Second Language Acquisition: Studies in Honor of Richard Schmidt*, Honolulu, HI: University of Hawaii, National Foreign Language Resource Center, 2013, p. 183-205.
- HIRSH, Eric Donald Jr., « Reading Comprehension Requires Knowledge of Words and the World », *American Educator* 27.1 (2003) : p. 10-29.
- HOLMQVIST, Kenneth, Marcus NYSTRÖM, Richard ANDERSSON, Richard DEWHURST, Halszka JARODZKA et Joost VAN DE WEIJER, *Eye Tracking: A Comprehensive Guide to Methods and Measures*, Oxford, New York : Oxford University Press, 2011.

- HUANG, Lingshan, Jinghui OUYANG, et Jingyang JIANG, « The Relationship of Word Processing with L2 Reading Comprehension and Working Memory: Insights from Eye-tracking », *Learning and Individual Differences*, 95 (2022): p. 102-143.
- HUANG, Ying 黄颖, « 印尼汉语学习者二语阅读中猜词策略调查与分析 » [Word-guessing strategies in L2 reading by Indonesian learners of Chinese: an investigation and analysis], Master Thesis, 厦门大学 [Xiamen University], 2008.
- HULSTIJN, Jan, « Incidental Learning in Second Language Acquisition », dans C. A. CHAPPELLE (éd.), *The Encyclopedia of Applied Linguistics. Vol. 5*, Wiley-Blackwell, 2013, p. 2632-2640.
- JEGERSKI, Jill, « Self-Paced Reading », dans J. JEGERSKI et B. VANPATTEN (éd.), *Research Methods in Second Language Psycholinguistics*, Routledge, 2014, p.20-49.
- JUST, Marcel A. et Patricia A. CARPENTER, « A Theory of Reading: From Eye Fixations to Comprehension », *Psychological Review*, 87.4 (1980): p.329-354.
- LAUFER, Batia et Jan HULSTIJN, « Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement », *Applied Linguistics*, 22.1 (2001): p. 1-26.
- LIU, Duo et Catherine MCBRIDE-CHANG, « Morphological Processing of Chinese Compounds from a Grammatical View », *Applied Psycholinguistics*, 31.4 (2010): p. 605-617.
- LIU, Zhifang 刘志方, Zhijun ZHANG 张智君, et Wei LIU 刘炜, « 阅读知觉广度内中文词汇的视觉信息提取：眼动研究证据 » [The visual coding of words in perceptual span in Chinese reading: evidences from eye movements], *应用心理报 [Chinese Journal of Applied Psychology]*, 21.4 (2015): p. 317-324.
- MALONE, Jonathan, « Incidental Vocabulary Learning in SLA: Effects of Frequency, Aural Enhancement, and Working Memory », *Studies in Second Language Acquisition*, 40.3 (2018): p. 651-675.
- NAGY, William E., Richard C. ANDERSON et Patricia A. HERMAN, « Learning Word Meanings from Context during Normal Reading », *American Educational Research Journal*, 24.2 (1987): p. 237-270.
- NAGY, William E., Patricia A. HERMAN et Richard C. ANDERSON, « Learning Words from Context », *Reading Research Quarterly*, 20.2 (1985): p. 233-253.
- NATION, I. Paul, « The Different Aspects of Vocabulary Knowledge », dans S. WEBB (éd.), *The Routledge Handbook of Vocabulary Studies*, Routledge, 2019, p. 15-29.
- PARIBAKHT, T. Sima, et Marjorie WESCHE, « Reading and “Incidental” L2 Vocabulary Acquisition: An Introspective Study of Lexical Inferencing », *Studies in Second Language Acquisition*, 21.2 (1999): p. 195-224.

- PARIBAKHT, Sima et Marie-Claude TREVILLE, « L'inférence lexicale chez des locuteurs de français et des locuteurs de persan lors de la lecture de textes anglais : Effet de la lexicalisation en première langue », *The Canadian Modern Language Review*, 63.3 (2007) : p. 399-428.
- PARIBAKHT, Sima et Marjorie WESCHE, « Reading Comprehension and Second Language Development in a Comprehension-based ESL Program », *TESL Canada Journal* (1993) : p. 9-29.
- PETERSON, Jeff Lynn, « Incidental Learning of Japanese through Reading Online, in Print, and in Digital Games », Mémoire du Master, Brigham Young University, 2016.
- PIGADA, Maria et Norbert SCHMITT, « Vocabulary Acquisition from Extensive Reading: A Case Study », *Reading in a Foreign Language*, 18.1 (2006) : p. 1-28.
- PULIDO, Diana, « The Effects of Topic Familiarity and Passage Sight Vocabulary on L2 Lexical Inferencing and Retention through Reading », *Applied Linguistics* 28.1 (2007) : 66-86.
- QIAN, Xujing 钱旭菁, « 词义猜测的过程和猜测所用的知识 — 伴随性词语学习的个案研究 » [Word meaning inferring and types of knowledge contributing to the process of inferring: a case study on incidental vocabulary acquisition], *世界汉语教学 [Chinese Teaching in the World]*, 1 (2005) : p. 87-96.
- RAYNER, Keith, Simon P. LIVERSEDGE, Sarah J. WHITE et Dorine VERGILINO-PEREZ, « Reading Disappearing Text : Cognitive Control of Eye Movements », *Psychological Science*, 14.4 (2003) : p. 385-388.
- SCHMIDT, Richard, « Consciousness and Foreign Language Learning: A Tutorial on the Role of Attention and Awareness in Learning », dans R. SCHMIDT (éd.), *Attention and Awareness in Foreign Language Learning (Teaching Report #9)*, Honolulu, HI : University of Hawaii, Second Language Teaching & Curriculum Center, 1995, p. 1-63.
- SCHMIDT, Richard, « Attention », dans P. ROBINSON (éd.), *Cognition and Second Language Instruction*, Cambridge : Cambridge University Press, 2001, p. 3-32.
- SOK, Sarah et Zhaohong HAN, « A Study of L2 Vocabulary Acquisition under Incidental and Intentional Conditions », *Vigo International Journal of Applied Linguistics* 17 (2020) : p. 113-140.
- VANPATTEN, Bill, « Evaluating the Role of Consciousness in Second Language Acquisition : Terms, Linguistic Features & Research Methodology », *AILA Review* 11 (1994) : p. 27-36.

- WANG, Wenjing 王文静, « 中文阅读过程中信息提取时间及词频效应的眼动研究 » [Eye movement study on information acquisition and frequency effect in Chinese reading], Master Thesis, 天津师范大学 [Tianjin Normal University], 2007.
- XIANG, Huiwen 向慧雯, Xinzhuo HU 胡欣灼, Runrong YANG 杨闰荣 et Xue SUI 隋雪, « 中文词切分的依据及眼动控制模型的解释 » [The basis of Chinese word segmentation and explanation of eye movement control model], 心理科学 [Journal of Psychological Science] 5 (2018): p. 1097-1102.
- XU, Jie 许洁 et Haolong WANG 王豪龙, « 阅读行为眼动跟踪研究综述 » [A review of studies on behavioral eye tracking in reading], Publishing Journal [出版科学] 28.2 (2020): p. 52-66.
- YANG, Cheng, Jiayi HAN, Rui SUN et Shiyi LU, « Predicting Language Proficiency in Chinese as a Second Language Using Eye Tracking », 2024: <https://ssrn.com/abstract=5003648> (consulté le 2 juin 2025).
- YU, Xi et Frank BOERS, « Inferring the Meaning of Idioms: Does Accuracy Matter for Retention in Memory? », *RELC Journal*, 2023.
- ZANG, Chuanli, Manman ZHANG, Xuejun BAI, Guoli YAN, Bernhard ANGELE et Simon P. LIVERSEDGE, « Skipping of the Very High Frequency Structural Particle de (的) in Chinese Reading », *Quarterly Journal of Experimental Psychology* (2017): p. 1-10.
- ZHANG, Haomin, « Concurrent and Longitudinal Effects of Morphological Awareness on Reading Comprehension among Chinese-speaking Children », *Read Psychol.*, 37 (2016): p. 867-884.
- ZHANG, Haomin et Keiko KODA. « Vocabulary Knowledge and Morphological Awareness in Chinese as a Heritage Language (CHL) Reading Comprehension Ability », *Read. Writ.*, 31 (2018): p. 53-74.
- ZHANG, Haomin, Xing ZHANG, Mengjie LI et Yiming ZHANG, « Morphological Awareness in L2 Chinese Reading Comprehension: Testing of Mediating Routes », *Frontiers in Psychology*, 12 (2021).
- ZHOU, Jing, « The Incidental Learning of L2 Chinese Vocabulary through Reading », *Reading in a Foreign Language*, 32.2 (2020): p. 169-193.

Auteurs

Anna ANASTASSIADIS-SYMÉONIDIS est Professeure émérite de linguistique à l'université Aristote de Thessalonique (Grèce). Elle a publié 10 ouvrages et plus de 200 articles dans des revues et des volumes collectifs. Elle a enseigné comme professeure invitée dans plusieurs universités françaises. Elle est spécialisée dans l'étude du lexique (néologie, emprunts, construction des mots, phraséologie, terminologie et enseignement du vocabulaire) ainsi que dans la rédaction de dictionnaires imprimés et électroniques du grec moderne.

Alberto BRAMATI est professeur associé de Langue et traduction française à l'università degli Studi di Milano (Italie). Traducteur de textes littéraires et d'essais en sciences humaines, ses recherches portent surtout sur les problèmes grammaticaux et rhétoriques de la traduction du français vers l'italien. Il vient de publier une grammaire contrastive spécifiquement conçue pour les besoins des traducteurs (*Grammatica contrastiva per i traduttori dal francese*, Libreria Cortina, Milano, 2025).

Serena DAL MASO est linguiste, docteure en linguistique et chercheuse en acquisition et didactique des langues au Département Cultures et Civilisations de l'université de Vérone (Italie). Ses recherches portent principalement sur l'acquisition de la langue académique chez des apprenants d'italien L2, en particulier chez des élèves issus de l'immigration, sur leurs compétences en lecture et en compréhension écrite et sur le rôle de la conscience morphologique et du traitement du lexique morphologiquement complexe.

Asimakis FLIATOURAS est maître de conférences en linguistique historique à l'université Démocrite de Thrace (Grèce). Il a publié plus de 80 articles dans des revues, des volumes collectifs et des actes de colloques et rédigé ou dirigé 10 ouvrages. Ses travaux portent sur les relations entre synchronie et diachronie, avec un accent particulier sur la langue grecque, notamment sur la morphologie diachronique, l'étymologie, les relations entre le grec et d'autres langues, l'onomastique et la terminologie.

Hélène GIRAUDO est chargée de recherche hors classe au CNRS et membre du laboratoire CLLE à l'université Toulouse – Jean-Jaurès. Spécialiste de psycholinguistique, elle étudie le traitement morphologique, la structuration du

lexique mental et l'acquisition des unités morphologiques en langue maternelle et seconde. Ses recherches, menées sur plusieurs langues dont le français, l'anglais, l'italien, l'espagnol, le grec et la langue des signes française, ont des applications en éducation, linguistique théorique et troubles du langage.

Jing GUO est maîtresse de conférences en chinois à l'Inalco (Institut national des langues et civilisations orientales) en France. Membre du laboratoire Plidam, ses recherches portent sur la didactique des langues étrangères, dans une perspective d'ingénierie éducative. Elle s'intéresse particulièrement au rôle des micro-compétences et des stratégies dans le développement des compétences orales. Elle pilote et analyse plusieurs projets pédagogiques, dont un programme eTandem et un atelier de théâtre en chinois.

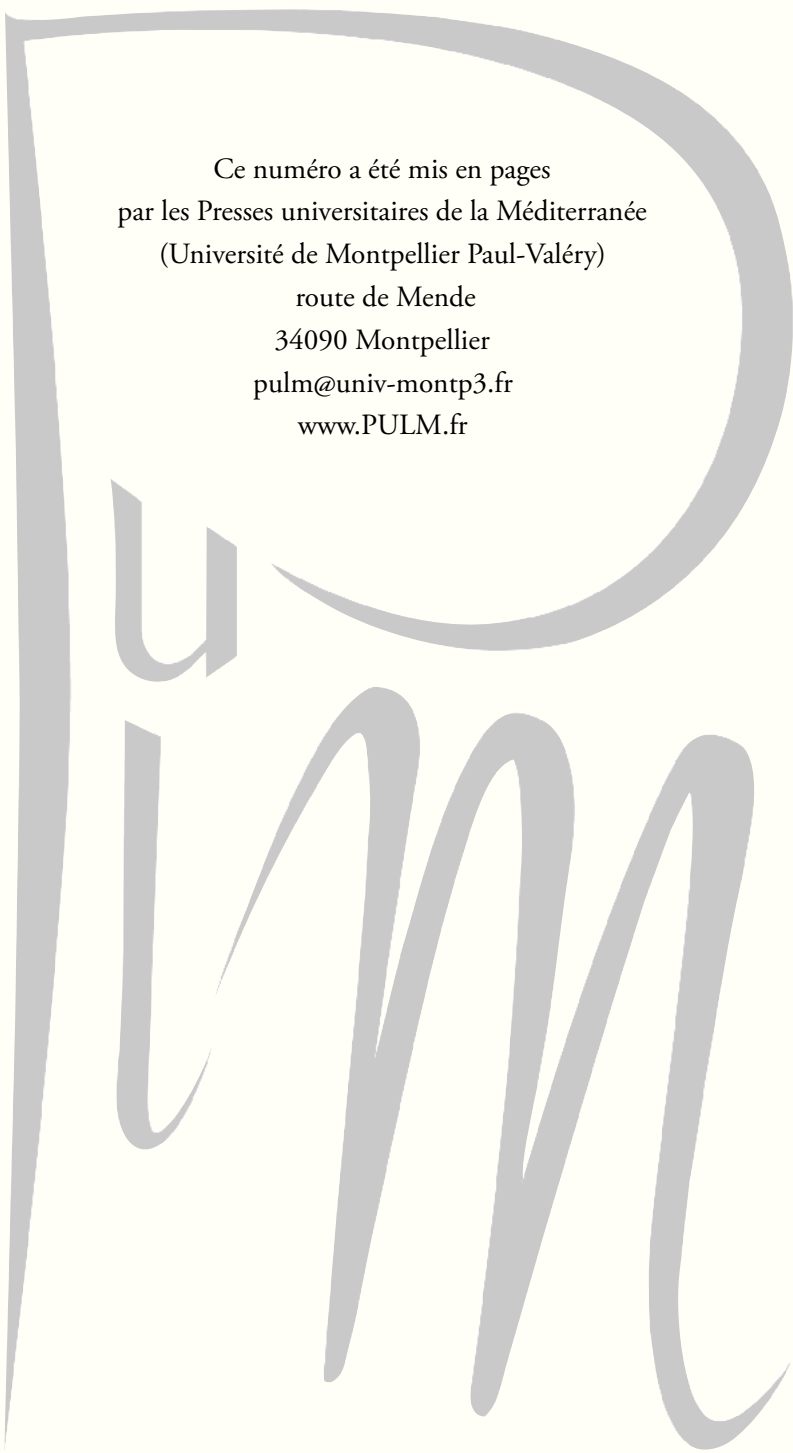
Enhao LÉGER-ZHENG est docteur en sciences du langage et membre du Laboratoire NeuroPsychoLinguistique (LNPL, UR 4156) de l'université de Toulouse — Jean-Jaurès. Ses recherches portent sur l'acquisition incidente du vocabulaire chinois L2 dans la compréhension écrite, en se penchant sur l'existence d'un équivalent d'une unité lexicale en L2 dans la langue maternelle de l'apprenant.

Xiaoliang LUO est maître de conférences à l'UFR LCAO, université Paris Cité. Il y enseigne la langue et la linguistique chinoises. Il est membre du laboratoire Histoire des Théories Linguistiques (UMR7597). Ses recherches portent principalement sur la phonologie et sur l'histoire de la linguistique.

Nikolaos NTAGKAS est chercheur postdoctoral à l'université Aristote de Thessalonique (Grèce). Il est titulaire d'un doctorat en linguistique théorique et appliquée, obtenu à cette même université ainsi que d'un master en linguistique historique et balkanique et d'une licence en philologie grecque. Ses recherches portent sur la morphologie et le traitement morphologique, la phonologie et la phonétique, la linguistique historique et les variétés du grec moderne, ainsi sur que la linguistique appliquée à la didactique.

Sabrina PICCININ est docteure en linguistique et chercheuse en acquisition et didactique des langues au Département de Langues et Littératures Étrangères de l'université de Vérone (Italie). Ses travaux portent sur le développement de la morphologie dérivationnelle, les processus de littératie et la compréhension des textes, ainsi que sur les compétences morphologiques des apprenants non natifs de l'italien et la conception d'interventions didactiques, dans une perspective fondée sur l'inclusivité et l'accessibilité.

Madeleine VOGA est Professeure à l'université de Montpellier Paul-Valéry, où elle assure la direction adjointe (2020-2025) de l'équipe ReSO, la direction du département de grec moderne, ainsi que celle de plusieurs thèses. Ses travaux associent l'étude linguistique, notamment lexicale, aux mesures chronométriques relatives à la perception du langage et les questions liées à la morphologie et sa représentation au traitement langagier monolingue et bilingue, ainsi qu'aux transferts inter-langues.



Ce numéro a été mis en pages
par les Presses universitaires de la Méditerranée
(Université de Montpellier Paul-Valéry)
route de Mende
34090 Montpellier
pulm@univ-montp3.fr
www.PULM.fr

