

Avis de Soutenance

Monsieur Ding YAN

Psychologie et ergonomie

Soutiendra publiquement ses travaux de thèse intitulés
*le role de la congruence orthographique dans l'acquisition de vocabulaire d'une langue étrangère -
quantifier la congruence par des statistiques infralexicales*

Travaux dirigés par Madame Séverine CASALIS et Monsieur Paolo MAIRANO

Soutenance prévue le **mercredi 08 juillet 2026** à 14h00

Lieu : Domaine Universitaire du Pont de Bois, 3 Rue du Barreau, Bâtiment F, 59650 Villeneuve-d'Ascq,
France
Salle : F0.44

Composition du jury proposé

Mme Séverine CASALIS	Professeure	Université de Lille	Directrice de thèse
M. Paolo MAIRANO	Professeur	Université de Tours	Co-directeur de thèse
Mme Eva COMMISSAIRE	Maître de conférences	Université de Strasbourg	Rapporteure
Mme Xenia SCHMALZ	Enseignant-Chercheur	Université de Munich	Rapporteure
M. Athanassios PROTOPAPAS	Professor	University of Oslo	Examineur
M. Arthur SAMUEL	Directeur de recherche	BCBL	Examineur
M. Walther VAN HEUVEN	Associate Professor	University of Nottingham	Examineur

Mots-clés : L2, Orthographe, Quantification

Résumé :

Tandis que le rôle de l'information orthographique pendant l'apprentissage est bien établi dans les études en langue première (L1), où un consensus suggère son effet facilitateur, les études en langue seconde (L2) sont loin d'arriver à un consensus similaire. La question de si oui ou non y a un effet de facilitation orthographique dans un contexte d'apprentissage en L2 est dépendant de nombreuses facteurs, dont un est basé sur le concept de la congruence orthographique. Même si des langues partagent la majorité de leur graphèmes, ce n'est pas nécessairement le cas pour leur correspondances-graphèmes-phonèmes (CGP) qui sont susceptibles de différer. Cette différence est pertinente puisque l'étude sur la congruence orthographique a démontré que lorsque des CGPs d'une L2 diffèrent des CPGs d'une L1, l'interférence introduite par cette incongruence peut négativement impacter des performances d'apprentissage en L2. L'approche dominante pour déterminer la congruence consiste dans des jugements émis par des experts en L2, qui catégorisent des mots en L2 selon leur expérience langagière: une catégorisation strictement dichotomique qui considère des mots en L2 comme étant soit congruent ou non. Malgré le fait que ce critère permet

de distinguer deux catégories qualitativement différentes, cette approche est fondamentalement subjective et restreint le concept de congruence sur seulement deux extrêmes. Cela lève un nombre de problèmes liés à la variabilité inhérente des tendances CGPs des langues, que cette approche ne peut pas prendre en compte. Pour aborder ces problèmes, nous proposons une reconceptualization de la congruence orthographique: au lieu de positionner la congruence comme deux extrêmes absolus, ces deux extrêmes peuvent être placés sur des cotés opposés d'un continuum. Le présent projet approche le concept de congruence sous un autre angle d'attaque pour mieux comprendre le concept ainsi que son rôle dans l'apprentissage de mots en L2. Étude 1 établit une quantification de la congruence orthographique en s'appuyant sur le concept des statistiques infralexicales. Cela est accompli grâce au G2P% framework, une solution logicielle qui attribue des probabilités d'occurrence graphème-phonème aux CGPs d'une L2 pour fournir une approximation de congruence pour n'importe quel mot en L2. Étude 2 examine la validité du G2P% framework, et confirme que l'utilisation des statistiques infralexicales d'une L1 reste viable dans un contexte explicitement L2, même si une L2 donnée est familière à une population en question; une validation qui ouvre une discussion sur le concept de la familiarité en L2. Même en absence de compétences langagières en L2, des adultes typiques ont vraisemblablement une familiarité avec un nombre de L2s, simplement du à une exposition générale au quotidien ce qui leur apporte un niveau de connaissances en L2 implicites, qui est rarement considéré, mais justifie une investigation plus approfondie. Étude 3 applique le G2P% framework à l'étude de la congruence en examinant les mécanismes qui soutiennent l'effet interférence d'incongruence (EII); et propose une formalisation de l'EII qui est structuré autour du concept des attentes CGPs. Nos résultats suggèrent que des apprenants L2 naïfs génèrent des attentes phonologiques lorsqu'ils sont exposés à l'orthographe d'une L2, mais plus intéressamment, qu'ils génèrent également des attentes orthographiques lorsqu'ils sont exposés à la phonologie d'une L2; des attentes qui sont fondamentales pour l'émergence des EII. En somme, trois études qui suggèrent que le concept de la congruence orthographique puisse être conceptualisé de manière continue, suggéré par des patterns de résultats linéaires et progressifs qui suivent le degré de congruence de mots en L2. Tandis que le G2P% framework est essentiellement un outil méthodologique, les théories qui le soutiennent nous permettent d'étudier la congruence orthographique sous un point de vue novateur.

Abstract:

Whereas the role of orthographic information during learning is firmly established in the first language (L1) vocabulary learning literature, which agrees that it has a facilitatory effect, the second language (L2) vocabulary learning literature is far from a similar agreement. Whether orthographic facilitation effects do or do not occur in an L2 learning context is dependent on a number of factors, one of which is grounded in the concept of orthographic congruence. Even when languages share most of their graphemes, the same does not necessarily apply to their grapheme-to-phoneme correspondences (GPC) which are susceptible to differ between languages. This difference is of consequence as the study of orthographic congruence has shown that whenever L2 GPCs differ from L1 GPC conventions, the ensuing incongruence introduces an interference that can negatively impact L2 vocabulary learning performances. The dominant approach for determining congruence consists in speaker-based judgments, which categorize L2 words on the basis of overall L2 knowledge: a strictly dichotomic categorization that considers L2 words as either congruent or not. Despite allowing for the meaningful differentiation between two categories, this approach is fundamentally subjective and constrains the concept of orthographic congruence to solely two extremes. This raises a number of issues, that are linked to the inherent variability of GPC tendencies across languages, and that this approach cannot account for. To address these issues, we propose a reconceptualization of orthographic congruence: instead of positioning orthographic congruence as two discrete extremes, those two extremes can be placed on opposite ends of a continuum. The present PhD project approaches the concept of orthographic congruence from

another angle of attack to further our understanding of the concept itself and its role in L2 vocabulary learning. Study I establishes a quantification of orthographic congruence that leans on the concept of infralexical statistics. This is accomplished through the G2P% framework, a software solution that attributes L1 grapheme-to-phoneme probabilities of occurrence to L2 GPCs in order to provide an approximation of orthographic congruence for any given L2 word. Study II examines the validity of the G2P% framework, and confirms that the use of L1 infralexical statistics remains viable in an overt L2 context, even when a given L2 of interest is familiar to a studied population; a validation that opens up the discussion on the concept of L2 familiarity. Even in the absence of any L2 proficiency, typical adults are likely to be familiar with a number of L2s simply due to everyday exposure which provides a level of implicit L2 knowledge that is seldom considered, but warrants further examination. Study III applies the G2P% framework to the study of orthographic congruence itself, by investigating the mechanisms that underlie the incongruence interference effect (IIE); and proposes a formalization of the IIE that is built around the concept of L2 GPC expectations. Findings suggest that naïve L2 learners generate phonological expectations when exposed to L2 orthography, but more interestingly, that they also generate orthographic expectations when exposed to L2 phonology; expectations that are fundamental for the emergence of IIEs. In sum, three studies that suggest that the concept of orthographic congruence can indeed be conceptualized as a continuum, as supported by highly linear and especially gradual patterns of results that are aligned with the degree of orthographic congruence of L2 words. Whereas the front-end of the G2P% framework consists of a methodological tool, the tool itself is upheld by a framework of underlying theories that open up the study of orthographic congruence to novel vistas of research that were not conceivable prior to its conception – vistas that are discovered and explored all throughout the present manuscript.