



Faculté de philosophie, arts et lettres

Commission doctorale du domaine CDL

Institut Langage & Communication – CENTAL



Thinking through gaming

Applying Concept-based instruction for French
pronominal verbs to Japanese university students

Membres du jury :

Thèse réalisée par

Julien Zakhia Doueih

Thomas François, promoteur

Silvia Lucchini, lectrice

Daichi Yamamoto, lecteur extérieur

Piet Desmet, lecteur extérieur

Serge Bibauw, président

en vue de l'obtention du grade de

Doctorat en langues, lettres et traductologie

Louvain-la-Neuve

Année académique 2024-2025

僕は夢を見る。

時々僕にはそれがただ一つの正しい行為であるように思える。

I dream. Sometimes I think that's the only right thing to do.

(村上春樹、スプートニクの恋人 - Haruki Murakami, Sputnik Sweetheart)

Julien Zakhia Doueïhi, 2024.

Cover picture generated by Gemini 1.5 Flash (imagen 3).

Games were created with RPG Maker MZ©, licensed version.

Analyses conducted in R (R Core Team, 2022), using *Metrics*, *readxl*, and many other packages. Plots generated in R, using packages *ggplot2* and *effects*.

Text written in British English.

Abstract

The present work attempts to introduce mobile-assisted language learning in the acquisition of French pronominal verbs for Japanese university students. Pronominal verbs are known as an heterogeneous category of verbs hampered by inconsistent teaching approaches and fragmentary descriptions of their uses. These difficulties are further increased when learners' first language lacks a corresponding structure. Japanese learners of French, in particular, are known to struggle with the acquisition of this grammatical concept, resulting in poor understanding and/or mastery of pronominal verbs up to upper-intermediate linguistic levels.

To compensate for these difficulties, the effects of two methodologies will be compared along with a control group. The first approach, inspired by the socio-constructivist theory of Concept-based Instruction, aims at depicting the target concept in a thorough fashion and promote its acquisition through verbalisations in learners' first language. The second approach, stemming from Communicative Language Teaching and Task-based Language Teaching, seeks to enhance conceptual understanding with contextualised output sequences and a non-linear learning structure.

Both approaches were integrated into mobile games so as to incentive learners' autonomy and implication, especially when considering the Japanese mobile game market. After being tested in a pilot experiment in October 2023, the games were used by 94 learners from 5 Japanese universities in May-June 2024. Each experiment was preceded by a pre-test and concluded with a post-test after a 3-week period during which players were free to play their assigned game at will. The results of the study, as demonstrated by both statistical and qualitative analyses, show that not only did each game contribute to significant improvements regarding the understanding of pronominalisation relative to task investment, they also brought different effects depending on test items. The success of the experiments calls for the implementation of mixed approaches in mobile-assisted language learning, taking the assets of both CBI and TBLT into powerful instruction tools for grammar acquisition.

Résumé

Cette étude a pour objectif d'introduire des outils d'apprentissage des langues assisté par des appareils mobiles (ALAM) dans l'acquisition des verbes pronominaux français à l'attention d'étudiants japonais à l'université. Les verbes pronominaux constituent une catégorie verbale hétérogène aux descriptions fragmentaires et à l'enseignement complexe. Ces difficultés sont d'autant plus grandes pour des apprenants dont la langue première ne possède pas de structure correspondante. Les apprenants japonais de français, en particulier, rencontrent régulièrement des problèmes dans l'acquisition de ce concept grammatical, comme démontré par une compréhension et/ou une maîtrise insuffisantes jusqu'à des niveaux linguistiques intermédiaires-avancés.

Au regard de ces difficultés, la présente étude aura pour objet la comparaison des effets de deux méthodologies en rapport avec un groupe contrôle. La première approche, inspirée des théories socio-constructivistes et nommée approche par concepts, vise à décrire un concept-cible d'une manière exhaustive et à encourager son acquisition par la verbalisation dans la langue première. La seconde approche, issue des approches communicative et par les tâches, cherche à améliorer la compréhension du concept en proposant des séquences de production contextualisées et un apprentissage focalisé sur la complétion de mini-jeux.

Les deux approches ont été implémentées dans des jeux pour téléphone mobile afin de promouvoir l'autonomie et l'implication des apprenants, notamment au regard du succès du marché du jeu mobile au Japon. À la suite d'une première expérience pilote en octobre 2023, les jeux ont été utilisés par 94 apprenants issus de 5 universités japonaises entre mai et juin 2024. Chaque expérience était précédée d'un pré-test et suivie d'un post-test réalisé après une période de jeu autonome de 3 semaines. Les résultats de l'étude montrent, statistiquement et qualitativement, que les deux jeux ont grandement contribué au développement de la conceptualisation des verbes pronominaux, avec des effets différenciés selon le traitement. Le succès de ces expériences montre l'intérêt de l'ALAM dans l'acquisition de la grammaire, ouvrant la voie à une approche mixte incluant conceptualisation et tâches.

Acknowledgments

Je voudrais remercier tous ceux qui ont aidé, de près ou de loin, à la réalisation de cette thèse de doctorat.

Merci à Thomas, qui m'a soutenu en tant que promoteur, mais également en tant que mentor. Tu m'as apporté la soif de rigueur et de perfectionnement, tout en m'offrant ton amitié et un soutien moral très important tout au long de ce travail. Merci aux membres du comité, Silvia et Daichi, pour avoir cru en moi initialement et accepté de suivre le bon déroulé de ma thèse jusqu'à son achèvement.

Merci aux membres du Cental, qui m'ont offert joie et paix lorsqu'il a fallu serrer les dents, mais aussi pour leurs conseils scientifiques, toujours à propos.

Merci à l'ambassade du Japon en Belgique, ainsi qu'à tous les membres du Ministère et de l'Université de Fukuoka qui m'ont permis de financer mes recherches pendant une longue période en montrant leur intérêt pour mon projet.

Merci au Paysan, qui m'a offert tant de joie de vivre en période de pandémie, alors que je vivais une période difficile. Merci à Sandro, qui a toujours été un exemple pour moi et un ami fidèle.

Merci du fond du coeur en particulier à mes parents, pour m'avoir offert un abri et pour avoir toujours soutenu mon entreprise, que cela soit moralement, financièrement et même parfois d'un oeil scientifique avisé. Je suis désolé d'avoir pris plus de temps que prévu pour achever ce travail.

Enfin, merci à mon amour, ma まな, soutien indéfectible qui n'a jamais cessé de croire en moi, malgré tous les doutes dont j'étais assailli tout au long de mon parcours de doctorant. Je n'aurais jamais pu aller aussi loin sans toi. Je t'aime.

Table of contents

List of Abbreviations.	12
List of Tables	13
List of Figures	14
Chapter 1: Introduction	
1.1. Research starting point and objectives	15
1.2. Common problems in the acquisition of French pronominal verbs	17
1.3. Context of foreign language acquisition in Japan	22
1.4. Challenges in the acquisition of pronominal verbs in Japanese universities	27
1.5. Structure of the thesis	33
Chapter 2: Theroretical framework	
2.1. Concept-based Instruction for applied L2 acquisition	37
2.1.1. Vygotsky's socio-cultural theory of mind	39
2.1.2. Concept-based instruction	41
2.1.3. Research methodology	44
2.1.4. Overview of studies carried out from 2003 to 2025	49
2.1.5. Challenges and avenues for future research in CBI.	58
2.1.6. Conclusion	63
2.2. Describing the target concept: pronominal-verbal constructions	64
2.2.1. What is a pronominal verb?	66
2.2.2. Review of the literature on classifications of pronominal verbs	68
2.2.3. Application of diathesis to pronominal verbs	75
2.2.3.1. About the concept of diathesis	75
2.2.3.2. Diathesis with pronominal verbs.	77
2.2.4. Two-step analysis	79
2.2.4.1. First phase:	
Identification of diatheses applicable to pronominal verbs.	79
2.2.4.2. Classification of pronominal verbs based on diatheses	83
2.2.4.3. Second phase:	
Validation through corpus analyses of Sequoia and FTB	86
2.2.5. Discussion about polyfunctional categories	90

2.2.6.	Insights and limitations	97
2.3.	Conclusion	101

Chapter 3: Research Methodology – System Design

3.1.	Digital gaming for language learning	107
3.1.1.	Learning through gaming: state of the art.	107
3.1.2.	Applying CBI to gaming	116
3.2.	The use of RPG Maker MZ as developping tool	118
3.2.1.	Presentation of RPG Maker	119
3.2.2.	RPG Maker Games design	120
3.2.2.1.	Deployment of the RPG Maker Games	121
3.2.2.2.	CBI-inspired variant	122
3.2.2.3.	CLT/TBLT-inspired variant	133
3.2.3.	Limits in game design choices	147

Chapter 4: Research Methodology – Experimental Design

4.1.	Participants	151
4.2.	Evaluation procedure	152
4.3.	Evaluation material	155
4.3.1.	Pre-/post-test metrics and design.	160
4.3.2.	Learner profile questionnaire and vocabulary size test	165
4.3.3.	Data collection	170
4.4.	Procedure	173
4.4.1.	Pilot experiment (2023)	173
4.4.2.	Final experiment (2024).	175
4.5.	Research questions	179
4.6.	Statistical methodology	183

Chapter 5: Experimental data analysis

5.1.	Descriptive statistics	187
5.2.	Linear mixed models	198

5.3.	Learners responses analysis	.	.	.	.	.	223
5.4.	General discussion	.	.	.	.	.	240
Chapter 6: Perspectives and conclusion							
6.1.	Summary of Findings	.	.	.	.	.	247
6.2.	Limits of present research	.	.	.	.	.	255
6.3.	Concluding notes	.	.	.	.	.	258
Bibliography		.	.	.	.	.	261
Appendix		.	.	.	.	.	281

List of Abbreviations

L1	First language
L2	Second language
FSL/FFL	French as Second Language/French as Foreign Language
ESL/EFL	English as Second Language/English as Foreign Language
CEFRL	Common European Framework of Reference for Languages
SLA	Second Language Acquisition
CL	Cognitive Linguistics
ZPD	Zone of Proximal Development
CB(L)I	Concept-Based (Language) Instruction
STI	Systemic Theoretical Instruction
CLT	Communicative Language Teaching
TBLT	Task-Based Language Teaching
TSLT	Task-Supported Language Teaching
DG(B)LL	Digital Game(-Based) Language Learning
(G)LMM	(Generalised) Linear Mixed Models
NPC	Non-Playable Character
QTE	Quick-Time Event
RPG	Role Playing Game
NLP	Natural Language Processing

List of Tables

<i>Table 1: Examples of common mistakes made by L2 French learners [...]</i>
<i>Table 2: Corpus annotation task results for Japanese learners</i>
<i>Table 3: Number of hits per keyword combination (restricted search)</i>
<i>Table 4: Studies distribution within the 5 databases</i>
<i>Table 5: Study context</i>
<i>Table 6: Study material</i>
<i>Table 7: Experimental design</i>
<i>Table 8: Fusion of middle and reflexive meanings (Schøsler, 1999)</i>
<i>Table 9: Valency reduction in pronominal sentences [...]</i>
<i>Table 10: Desclés and Guentchéva's classification (1993)</i>
<i>Table 11: Bień's (2007) classification</i>
<i>Table 12: Change from neutral to passive (agentless) diathesis</i>
<i>Table 13a & 13b: Main syntactic functions in French, Main semantic roles [...]</i>
<i>Table 14: Major diatheses in French</i>
<i>Table 15: Example of diathetic mapping</i>
<i>Table 16: Typology of diatheses represented by pronominal verbs</i>
<i>Table 17: Distribution of pronominal structures in FTB, comparison [...]</i>
<i>Table 18: Confusion matrix recall and true positive rates compared with [...]</i>
<i>Table 19: Overall comparison of our CBI and TBLT games</i>
<i>Table 20: Detailed content of the CBI-inspired game</i>
<i>Table 21: List of extracted in-game variables (CBI-inspired game)</i>
<i>Table 22: Integration of game assets into TBLT phases (Ellis, 2000)</i>
<i>Table 23: Difficulty degree differences (final task)</i>
<i>Table 24: Detailed content of the TBLT-inspired game</i>
<i>Table 25: In-game task types in the TBLT game</i>
<i>Table 26: List of extracted in-game variables (TBLT-inspired game)</i>
<i>Table 27: List of extracted variables</i>
<i>Table 28: Comparison of pre- and post-test score changes in the pilot [...]</i>
<i>Table 29: Overview of the conduct of the final experiment (2024)</i>
<i>Table 30: Pre-test scores by university and by exercise type</i>
<i>Table 31a: Learners' sex distribution</i>
<i>Table 31b: Learners' age distribution</i>
<i>Table 31c: Learners' class distribution</i>
<i>Table 31d: Learners' preliminary attitudes and perceptions by university [...]</i>
<i>Table 32: Learners' game time distribution</i>
<i>Table 33: Learners' game completion distribution</i>
<i>Table 34: Post-test scores by university and by exercise type</i>
<i>Table 35: Learners' perceived ease of use and usefulness by university [...]</i>
<i>Table 36: Pre- and post-test means evolution by treatment</i>
<i>Table 37: Linear Mixed Modelling for RQ1 (selected effects)</i>
<i>Table 38: Linear Mixed Effect Modelling (RQ2, E1)</i>

Table 39: Linear Mixed Effect Modelling (RQ2, E2)
 Table 40: Linear Mixed Effect Modelling (RQ2, E3)
 Table 41: Linear Mixed Effect Modelling (RQ2, E4)
 Table 42: Linear Mixed Effect Modelling (RQ2, E5)
 Table 43: Linear Mixed Effect Modelling (RQ3)
 Table 44: Evolution in pronominalisation correctness rate, E1
 Table 45: Evolution in pronominalisation correctness rate, E2
 Table 46: Evolution in pronominalisation correctness rate, E3
 Table 47: Evolution in pronominalisation correctness rate, E4
 Table 48: Evolution in pronominalisation correctness rate, E5
 Table 49: Global gains in pronominalisation correctness rate by diathesis

List of Figures

Figure 1: An example of yakudoku activity (英単語学習の科学)
 Figure 2: Example of yakudoku-inspired activity in French textbook *Moi, je...*
 Figure 3: Explanation of pronominal verbs in *Spirale, Nouvelle Édition* (p. 109)
 Figure 4: Exercises about pronominal verbs in *Café Français Nouveau*
 Figure 5: Architecture of this thesis
 Figure 6a & 6b & 6c: Diathesis distribution in *Sequoia*/FTB/combined corpora
 Figure 7: Flowchart representation of our 8-type taxonomy of pronominal verbs
 Figure 8: Games like *World of Warcraft* promote cooperation as a way to [...]

Figure 9: "Dream world" sequences allow the player to move around [...]

Figure 10: Example of SCOPA realisation

Figure 11: Optional verbalisation-type tasks are available once the main [...]

Figure 12: Branching dialogues also appear in interludes

Figure 13: The final task is modulable and decided from the start

Figure 14a & 14b: Optional theory lessons in the theater mix in...

Figure 15: Quizzes are optional activities used to practise [...]

Figure 16: Final task battle scene

Figure 17: Glossary entries fill as the player progresses through [...]

Figure 18: Experimental process

Figure 19: Relationship between abilities, skills and concept

Figure 20: An example of vocabulary size test item with open answer

Figure 21: QR codes were scanned by learners to pass the pre-test and the [...]

Figure 22: Results of vocabulary size test, sample by university

Figure 23: Relationship between Pre-test score and Pre-post score change [...]

Figure 24: Impact of Time on Task on learning gains (RQ1)

Figure 25: Impact of interacting Game metrics on learning gains (RQ1)

Figure 26: Impact of Time on Task on E1 learning gains (RQ2)

Figure 27: Impact of Time on Task on E2 learning gains (RQ2)

Figure 28: Impact of interacting Game metrics on E3 learning gains (RQ2)

Figure 29: Impact of Time on Task: Completion on E4 learning gains (RQ2)

Figure 30: Impact of Time on Task: Completion on E5 learning gains (RQ2)

Figure 31: Impact of Perceived ease of use on learning gains (RQ3)

Chapter 1: Introduction

1.1. Research starting point and objectives

At all times, linking experimental research with practical uses in foreign language instruction has been a persistent challenge for linguists. This is especially true regarding conceptual knowledge such as grammar. In order to teach a specific grammatical concept, L2 instructors need not only a thorough understanding of the concept itself, but also weigh how to balance instruction so as to leave enough practice space to learners while providing explanations. In that regard, the role of the L1 on the acquisition of L2 concepts represents an essential key to the puzzle. Indeed, determining whether L2 concepts should be taught through an explicit L1 module or in an implicit, L2-only module is not an easy task. Several studies (Butzkamm & Caldwell, 2009; de la Fuente, 2015; Hall & Cook, 2012) have shown that L2 acquisition should take into account the role of learners' L1, while others demonstrate negative transfer effects on the process of L2 learning (Erdocia & Laka, 2018; Sabbah, 2015). The divergence of results in this field of research has brought us to delve into the role and the impact of methodologies using and/or limiting the exposure to a L1 to teach a L2. However, rather than focusing on the acquisition of lexical knowledge and/or L1 interference on L2 performance, our efforts were concentrated in a less commonly treated question, that is to say the way how mental constructs in learners' L1 interfere with the acquisition of grammatical concepts in L2, especially in the context of autonomous learning with mobile applications.

Grammatical knowledge can be complex to teach to L2 learners, especially when the latter have no common ground of comparison within their mother tongue. The lack of reference in one's L1 is often considered an obstacle in the acquisition process since, unlike with the L1, L2 concepts are not salient

in the input, nor rooted into an individual's universal grammar (Brown, 1998; de la Fuente, 2015). As such, new grammar subjects taught in formal instruction must first be conceptualised to be understood and used in a fluent manner. The way L2 concepts are conceptualised when no corresponding structure exists in a learner's mother tongue can vary (Ortega, 2009), but it is undeniable that the more complex a concept is, the more natural it seems to use explicit L1 instruction to teach it, because learners themselves will use their L1 to feel comfortable with the new concept (Renoud, 2016). According to Swain & Lapkin (2013), the act of *linguaging* – using language to develop, organise and control a cognitive load – is necessarily made in L1 as the load becomes heavier, while easier concepts which have more transfer potential can be acquired with the sole use of the L2. Yet, as much as the role of these mental activities in the process of acquiring a L2 is usually agreed upon, the impact explicit L1 teaching has on the development of L2 conceptual knowledge is rarely explored, especially when compared to immersive and L2-heavy contexts.

The importance of *linguaging* in the process of L2 grammar acquisition cannot be undermined, especially in the event where the targeted learners do not have a corresponding structure in their L1. To address these challenges, we wished to provide tools to teach an exemplary – yet difficult to acquire – concept in French, select a target population without prior mastery of the concept or reference in their L1, then test the efficiency of our modules. After creating two mobile applications based on the same architecture but with different theoretical foundations, we decided to carry out a comparative experiment related to the acquisition of French pronominal verbs among Japanese L1 speakers. Pronominal verbs are known to be a productive concept in French, as many basic structures learned by novice learners

involve pronominalisation, such as *Je m'appelle...* ([My name is...]) or *Je me promène* ([I take a stroll]). Yet, without proper scaffolding, the rationale behind the use of pronominal verbs can be difficult to grasp for non-native speakers, even at intermediate to upper-intermediate levels of mastery. In turn, studies have shown that increasing the amount of *languageing* can help in alleviating the cognitive load of grammatical concepts supposedly mastered at later stages of the acquisition process (Fehrmann, 2016; Lai, 2012), enhancing the fact that pronominal verbs represented the ideal target for the application of teaching methods focused on increased conceptualisation and L1 scaffolding.

An overview of the numerous difficulties regarding the acquisition of French pronominal verbs will further be detailed in the following sections of this introductory chapter. Notably, we will portray the problems encountered by Japanese learners, among others, whenever attempting to use pronominal verbs, before showing that the way this concept is taught in formal instruction only exacerbates an already complex linguistic phenomenon.

1.2. Common problems in the acquisition of French pronominal verbs

Pronominal verbs are a central concept in modern French, as almost any transitive verb can be used in its pronominal form to hold some sort of reflexive meaning. However, many verbs also take different meanings depending on the context, the structure and the original verb, leading to a lot of variation in semantics (Burston, 1979). The heterogeneity of pronominal verbs causes several mistakes among L2 learners of French, as it is shown in error analyses. Duée (2002) notices that L1 Spanish speakers taking French classes tend to overuse pronominal structures when it is not needed, notably due to the influence of their mother tongue – in which pronominal verbs are

even more common. In specific structures however, such as passive-meaning, they would not always identify the necessity to use pronominal verbs. Likewise, Alsadag (2021) reports in a study with L1 Arabic students that they would often try to remember pronominal verbs without understanding the underlying meaning of the reflexive pronoun. This phenomenon is also commonly observed among Japanese learners, as they do not have any corresponding structure either in their L1 (Kambaru, 2023; Soga, 2003). Learners would then tend to try to avoid using these verbs as much as possible unless they could not find a substitute; for instance, many Japanese learners would rather use nominal structures such as *J'ai fait une promenade* than corresponding pronominal verbs *Je me suis promené* (Soga, 2003). While both sentences are correct and commonly used in French, the use of substitution strategies within learners who do not have pronominal verbs in their linguistic system *per se* can also lead to slower polishing of their understanding of this grammatical concept. As a consequence, French pronominal verbs can be more or less difficult to acquire depending on learners' first language as well as their overall mastery of French. A non-exhaustive list of examples of errors made by learners when using pronominal verbs can be found in Table 1:

<i>Reference</i>	<i>Example</i>	<i>Correction</i>	<i>Type of mistake</i>
Afifi (2018)	<i>Il était se laver.</i>	<i>Il se lavait.</i>	Incorrect form
Alsadag (2021)	<i>Je vais se laver.</i>	<i>Je vais me laver</i>	Incorrect agreement
Duée (2002)	<i>Il se lui reprochait.</i>	<i>Il lui reprochait</i>	Hypercorrection
Duée (2002)	<i>La tourmente qui se prolonge.</i>	<i>La tourmente qui se prolonge</i>	Omission
Karan (2016)	<i>Nous ne sommes pas encore préparés.</i>	<i>Nous ne nous sommes pas encore préparés.</i>	Omission
Lapkin & Swain (2000)	<i>Quelqu'une se dorme à l'envers dans le lit</i>	<i>Quelqu'un dorme à l'envers dans le lit</i>	Hypercorrection
Nguyen (2004)	<i>Souviens-tu Ba et Nam?</i>	<i>Te souviens-tu de Ba et Nam?</i>	Omission
Nguyen (2004)	<i>Ne inquiétez pas trop.</i>	<i>Ne vous inquiétez pas trop.</i>	Omission
Nguyen (2004)	<i>Thanh m'a invité à se promener avec lui.</i>	<i>Thanh m'a invité à me promener avec lui.</i>	Incorrect agreement
Nguyen (2004)	<i>Est-ce que tu veux se promener?</i>	<i>Est-ce que tu veux te promener?</i>	Incorrect agreement
Wakely (1998)	<i>Le problème qui s'apparaît</i>	<i>Le problème qui apparaît</i>	Hypercorrection

Table 1: Examples of common mistakes made by L2 French learners with pronominal verbs

As we can see in Table 1, mistakes in the use of pronominal verbs are common and take diverse forms: hypercorrections, omissions of pronouns as well as confusion in forms. The proportion of errors as well as their type depends on various factors. In a study involving four classes of Canadian immersion middle-schoolers, Lapkin & Swain (2000) reported a high number of mistakes in the use of pronominal verbs (ratio of correct pronominals to pronominals forms ranging from 0.23 to 0.35). While two classes used a sketch jigsaw as a basis to write a narrative, the others used a dictogloss (group-based summarization of L2 content). The researchers noticed that the class using a jigsaw tended to overuse pronominal verbs, while the others were more reluctant, most likely due to the lack of word input. These observations show that due to the lack of correspondence in

their L1 (in this case, English), it is not intuitive for FSL learners to determine in which contexts pronominal verbs should be used (Lapkin & Swain, 2000).

In many cases, mistakes in the use of pronominal verbs can be overcome by sheer practice and increased effort in learning; however, comprehension errors are much more problematic, as they might come from a lack of understanding of the concept itself (Swain et al., 2009). The influence of learners' L1 cannot be undermined either, as conceptual mistakes are more common when learners don't have a reference in their mother tongue. Nguyen (2004) reports how L1 Vietnamese learners had trouble figuring out the uses of pronominal verbs. Their mistakes would also come from a tendency to remember the verbs by using "stocked" knowledge rather than applying a pattern. As such, interlinguistic and intralinguistic influences can both have a part in the acquisition of pronominal verbs.

Another major source of problem concerning the acquisition of pronominal verbs for L2 learners comes from the limits in descriptions of these verbs. According to Wakely (1998), while any transitive verb can have a potential pronominal form, it is difficult if not impossible to predict which intransitive verbs will be reflexive and which will not; the division is only explained by "*careful historical analysis*" (Wakely, 1998: 423). Consequently, the descriptions in teaching of French pronominal verbs almost systematically include **subjective** or **lexicalised** categories, which encompass verbs without regular patterns or distinctive semantics. While there were attempts to categorise this type of verbs further with more thorough descriptions (Barque et al., 2018; Desclés & Guentchéva, 1993; Petersilka, 2012), the choice of categorisation criteria was not always consistent, as they would

sometimes rely on syntactic patterns, sometimes on semantic ones (Bieñ, 2007). This doesn't help clarify the otherwise already complex phenomenon of French pronominal verbs (Connors & Ouellette, 1996), nor does it allow instructors to rely on a satisfying method to teach these verbs (Delbarre, 2010; Kraus, 1997; Nakamura, 2004; Soga, 2003). Indeed, Allam (2010) brought an overview of how pronominal verbs are presented in CLT textbooks issued between 1982 and 2004 for French as second language (FSL). In *Archipel* and *Libre Echange*, the focus is put on formal properties : construction, past participle agreements then meaning. Even then, neither lexicalised verbs and benefactive pronominals (a combination of a reflexive verb + direct object) are mentioned. The authors also emphasise the role of pronominal verbs in expressing daily habits. The focus on routines is also present in *Reflets*, *Campus* and *Espaces*. *Reflets* is also the only textbook including meaning-focused exercises. On the other hand, all of the textbooks make a distinction between reflexive-meaning and reciprocal-meaning verbs, apart from *Tempo* and *Studio 60*. Lexicalised pronominal verbs are absent from all of the textbooks but *Cours de langue et de civilisation françaises*. As a matter of fact, every textbook follows its own agenda for the teaching of pronominal verbs. This lack of consistency is yet another major challenge which has to be considered when seeking to offer adequate teaching.

To summarise, the acquisition of pronominal verbs appears to be mostly hindered by the following factors: 1) heterogeneity of the grammatical class of pronominal verbs; 2) lack of reference/corresponding structures in learners' L1; 3) lack of consistency in grammatical descriptions and teaching methods.

We will see in the following sections that these factors are only exacerbated in the context of FSL in Japanese universities.

1.3. Context of foreign language acquisition in Japan

Japan is a particularly interesting case for research in SLA, as the country still represents, to this day, a linguistically homogenous environment. Indeed, according to Apple et al. (2013), 90% of Japanese learners exclusively study English as a foreign language. Even then, and despite having studied English for six or more years in middle school and high school when applying to a university, most Japanese still lack the confidence to use English to communicate (Dickinson, 2010). The instructional module in Japanese EFL classes, often non-communicative and form-focused, also has a negative impact on learners' motivation (Kikuchi & Sakai, 2009). One problem comes from the fact that the system of entrance exams in Japan puts the focus on normative knowledge instead of communicative skills, further delaying or putting a stop to the acquisition of English once the exams were passed (Butler & Iino, 2005). Students are also rarely motivated to use English outside of the classroom, mostly because of the low exposure rate to English in Japan (Chu & Nakamura, 2010).

The context of French in Japanese universities is not much different. According to Ishikawa (2011), foreign language teaching in Japan is mostly motivated by economic reasons – such as increasing one's chances to get a job – and, consequently, the choice of French as foreign language is becoming less and less obvious, as it has lost part of its influence compared to other languages like German or Chinese in the employment market. The poor recognition of French is also worsened by the fact that Japanese universities enforce English as a mandatory L2 since 1991, diverting students from exploring other languages (Mayaux, 2007). Furthermore, the way French is taught has barely changed since the 70s (Roussel, 2006), a time when French was considered as a counterpower to English in Japan and

was a popular language to learn (Kessler & Siary, 2008).

While the diminishing popularity of French in Japan is a concern for teachers, the methodology also brings its share of difficulties. In Japan, language learning is closely connected to a method called **yakudoku** (訳読, lit. “translation-reading”). This method consists of word by word translation, exclusive use of learners' L1 for grammatical explanations, as well as minor to no emphasis on class interaction (Nishino, 2008). According to Nishino (2008), about 70-80% of teachers are reported to use *yakudoku* to teach English. An example of yakudoku activity can be seen in Figure 1:

意味順☆訳読シート	Lesson 7D Never Despair
3年 組 番 名前	

[誰が | する・です | だれ | なに | どこ | いつ]

I've sometimes felt down, but I've never despaired.
 | | | |

I'm young and strong.
 | | |

I'm living through a big adventure.
 | | |

If I complain all the time, I won't have any fun.
 | | | | |

I'm happy and cheerful.
 | | |

Figure 1: An example of yakudoku activity (英単語学習の科学, p.18)

Historically, *yakudoku* was intensively used to translate foreign texts when Japan opened up to the West in the 19th century (Chevalier, 2008). This

method allowed learners to quickly access the meaning of a text without the need to acquire other communication skills, as it relies heavily on repetition and memorisation of a translated text (Chevalier, 2008; Noda & O'Regan, 2020). Conceptually speaking, *yakudoku* implies that the target content is primarily understood in Japanese, with no regard to L2 competence (Gorsuch, 1998: 9). As of today, it is still a popular choice for language teaching, even if communicative language teaching (CLT) is slowly taking over in universities – in 1997, CLT was reported to be used in 28% of Japanese universities, albeit it has been following an upward trend (Kambaru, 2023; Nowlan & Samuell, 2020; Suzuki, 2001).

The success of CLT methodologies for L2 in Japan is also hampered by the cultural impact of *washback* effect (O'Donnell, 2003). This phenomenon refers to the adaptation of L2 learning programs to test completion requirements. In Japan especially, entrance exams are notorious gatekeepers to students' future. As such, language classes are focused on delivering memorizable and translatable written content, which facilitates self-learning (Chevalier, 2008). Consequently, L2 instruction in Japanese classrooms is mostly focused on translation activities stemmed from the *yakudoku* tradition, and gives little to no space for communication skills to develop, even though students themselves generally express the wish to be granted more opportunities to practise their L2 (Ishikawa, 2011). This learning tradition also has the downside of putting many Japanese students in uncomfortable situations in the event where they are potentially confronted with communicative language teaching in higher education (Apple et al., 2013). This might also explain why even modern textbooks for French used in Japanese universities still choose to revolve around *yakudoku*-inspired activities (albeit to a lesser extent), as showed in Figure 2:

2 単数形 / 複数形の使い分け

あるものを一般的・抽象的・概念的な意味で話す場合、数えられる名詞には定冠詞を「複数形」で使い、数えられない名詞の場合は「単数形」で使います。

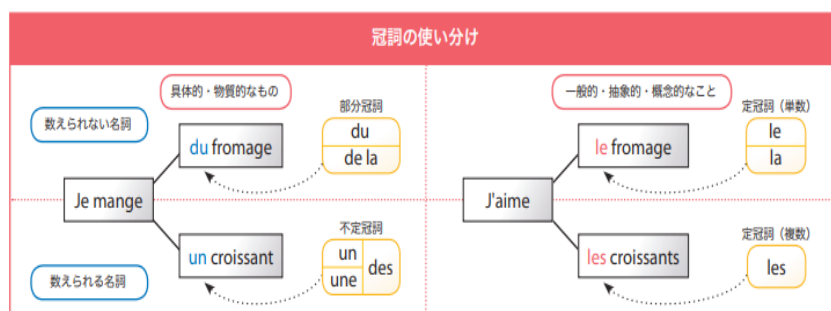


Figure 2: Example of yakudoku-inspired activity in French textbook *Moi, je...* (Miki et al., 2012: p.61)

Meanwhile, enforcing CLT for this peculiar context does not necessarily mean that Japanese students will become more proficient in their L2. In that regard, Kambaru (2023) reports that while Japanese universities slowly started to shift from a traditional grammar-focused approach towards communicative language teaching, it does not necessarily correlate with better language proficiency, notably for English. Indeed, Griggs (2007: 176) reminds us that “*interaction does not guarantee acquisition*”. As mentioned by Suzuki (2003) and Irie (2003), Japanese learners are very self-conscious about their image, as well as their own perception of their L2 skills. In the event that learners feel as though their language proficiency is lacking, which is common in Japanese schools, these negative perceptions often lead to low self-confidence and hesitancy in communicating (Kambaru, 2023). Consequently, attempts of CLT classrooms do not always show the expected results, as students would stumble when required to speak up and shy away from communication tasks. This is especially true when learners are asked to communicate in languages they have had no prior contact with before

attending university, such as French (Ishikawa, 2011). Furthermore, the “real-world” interactions promoted by CLT as means to enhance communicative skills are difficult to implement in a monolingual environment such as Japan (Nowlan & Samuell, 2020).

Another problem surrounding the application of CLT in Japanese classes is related to the fact that many Japanese students feel that they are not in a learning process when not explicitly taught grammar or vocabulary (Burrows, 2008). It also means that an explicit method using their L1 as a scaffolding tool could help them feel more comfortable with the learning process, rather than enforcing a L2-exclusive method - which is also unrealistic in the context of Japanese education, in which learners make extensive use of their L1 for languaging purposes anyway (Aoyama, 2020). According to Soga (2003), Japanese students often lack comprehension of complex grammatical matters such as aspect or pronominal verbs. Rather than putting the focus on the translation of literature or scientific texts, Soga believes that more effort should be put into making learners consider the productivity of each and every concept in the L2. Nakamura (1996; 2004) also reports that students tend to avoid using a lot of verbs in French because they do not understand their communicative uses. The hardships are even bigger in case of pronominal verbs, which have an additional layer of complexity. The researcher also points out that in most cases, conceptualisation of these verbs is in fact lacking. In his dissertation study on French verbal aspect learning among Japanese university students, Renoud (2016) notices how much cognitive work takes an important part in the acquisition of the target concept, deploring the lack of conceptualisation in textbooks traditionally used in Japan such as *Le français* or *Café Français*. Notably, Renoud asserts that while a communicative tasks should

be generally favoured in L2 learning, the role of the first language in the acquisition of a complex notion such as aspect cannot be undermined either.

Observations made throughout this section allow us to fathom that the frequent difficulties in L2 concept acquisition demonstrate a need for different and/or improved methods. The impact of the tradition of *yakudoku* in language teaching, as well as the lack of adaptation of CLT further impede the overall lack of exposure to foreign languages in Japan.

As the guiding line of our work, we chose to focus our research on the acquisition of French pronominal verbs among Japanese university students. The next section will show that the teaching of a concept such as pronominal verbs in Japan in fact brings its own challenges in language classes.

1.4. Challenges in the acquisition of pronominal verbs in Japanese universities

Pronominal verbs embody most of the problems which should be taken into consideration when teaching L2 concepts to Japanese learners. The lack of reference structure in Japanese and the intrinsic difference between the two verbal systems – unlike French, the Japanese language produces intransitive verbs much more - represent a challenge for Japanese learners (Nakamura, 2004). Consequently, this grammatical category is often put aside by Japanese learners, who would rather avoid using them by finding substitutes. In Japan especially, pronominal verbs are known to be insufficiently studied, as grammar manuals tend to introduce them out of context, in isolated sentences or through word-to-word translations (Delbarre, 2010). Indeed, instances of the way pronominal verbs are presented in textbooks commonly used in higher education French classes in Japan at the beginner level (A1/A2), such as in *Spirale Nouvelle Édition*, *Café Français Nouveau*,

Communication progressive du français or *Alter Ego 1*¹ all demonstrate that the underlying mechanics of this grammatical concept are often overlooked:

-The textbook *Spirale Nouvelle Édition* is composed of 21 lessons divided into 7 units (along with a few reminders of some grammatical concepts at the end of the textbook). Lesson 16 is partially focused on the teaching of pronominal verbs. A short summary (Figure 3) explains what a pronominal verb is (in fact, only the direct reflexive meaning of pronominal verbs is presented), shows conjugation forms for the *présent* and *passé composé* tenses and gives some information about formal properties (subject-verb agreement, auxiliary choice in past tenses).

5 La conjugaison des verbes pronominaux : s'appeler, se lever, se coucher, se promener...
 代名動詞の活用：名前は～です、起きる、寝る、散歩をする、など

➤ Un verbe pronominal est un verbe précédé d'un pronom personnel qui représente la même personne ou la même chose que le sujet. Le sujet exerce l'action sur lui-même.

➤ 代名動詞は、主語と同じ人物をあらわす人称代名詞(再帰代名詞といいます)を動詞の前に先立てる動詞です。

Présent de l'indicatif 直説法現在形			Passé composé 直説法複合過去形			
Je	me	promène	Je	me	suis	promené(e)
Tu	te	promènes	Tu	t'	es	promené(e)
Il/Elle/On	se	promène	Il/Elle/On	s'	est	promené(e)(s)
Nous	nous	promenons	Nous	nous	sommes	promené(e)s
Vous	vous	promenez	Vous	vous	êtes	promené(e)(s)
Ils/Elles	se	promènent	Ils/Elles	se	sont	promené(e)s

➤ Les verbes pronominaux forment toujours leur passé composé avec l'auxiliaire *être*.
 ➤ Leur participe passé s'accorde donc en genre et en nombre avec le pronom réfléchi (qui reprend le sujet qui parle).
 ➤ 代名動詞を複合過去形にするときは、つねに *être* を助動詞にします。
 ➤ その過去分詞は、話し手の主語と同じ人物をあらわす再帰代名詞と性・数が一致します。

Figure 3: Explanation of pronominal verbs in *Spirale, Nouvelle Édition* (p. 109)

While immediately accessible due to explanations in both French and Japanese, this description appears to be solely focused on how to form pronominal verbs; meanwhile, the explanations towards the meaning are undoubtedly fragmentary (no explanation whatsoever about what other

1 Café Français Nouveau, 2016: Gaillard, Kato, Nakagawa, Yoko Sudre, Yanagishima

meanings a pronominal verb can take, e.g. reciprocal, passive, etc.). Further looking into the lesson shows that pronominal verbs are presented for their uses in daily habits (such as *Je me lève à...*). The fact that the authors contextualised this concept by emphasizing the lexical field of daily lives should give an idea of one application of pronominal verbs to learners.

-*Café Français Nouveau*, a textbook including 15 lessons and an extensive exercise section shows a few exercises focused on pronominal verbs in Lesson 11, as can be seen in Figure 4. They allow learners to practise the formation of pronominal verbs by conjugating in the correct form and making the correct subject-verb agreement. As the abilities required here are exclusively put on knowledge of forms, we would not expect learners to have a solid conceptual understanding of the use of pronominal verbs at this point in the textbook.

-
3. 〈 se promener 〉を活用させましょう。
- | | |
|-----------------|-------------------|
| je _____ | nous _____ |
| tu _____ | vous _____ |
| il / elle _____ | ils / elles _____ |
4. 動詞を適切な形に活用させて、文を完成させましょう。
- 1) se promener Je _____ avec des amis.
 - 2) s'habiller Elle _____ pour une soirée.
 - 3) se lever Nous _____ à six heures moins le quart.
 - 4) s'appeler Vous _____ comment ?

Figure 4: Exercises about pronominal verbs in Café Français Nouveau

-In *Communication progressive du français*, Lesson 21 introduces pronominal verbs at the present tense. In this lesson, two meanings of pronominal verbs are related to: direct reflexive (*Je m'appelle...*) and reciprocal (*Vous vous connaissez?*). The focus is put on the formation of

these verbs, as there is no details about the relationship between the pronominal verb and its non-pronominal variant. In Lesson 24 and Lesson 29, a few other verbs are introduced as well in dialogues: *s'occuper*, *se mettre*, *se retrouver*, *se cogner contre*, *se tordre la cheville*, *se casser la jambe*. While the inclusion of these verbs shows the variety of meanings pronominal verbs can take, there is no explanation either about the reasons why they are used in their pronominal form.

-Finally, in *Alter Ego I*, pronominal verbs are described in Unit 4 (out of 10 Units). Like in *Spirale Nouvelle Édition*, they are immediately associated with daily habits (*Je me réveille à 6h30*, *Je m'endors*, *Les enfants se couchent tôt*, etc.). In that Unit, before the introduction to the conjugation table of *se réveiller* at the présent tense, learners are asked about the specificities of these constructions compared to other verbs. This is an interesting way of introducing the concept of pronominal verbs while not delving too deep into their potential meanings.

Through these depictions of pronominal verbs made in several textbooks used in Japanese universities, we can make a few observations:

- Pronominal verbs are commonly introduced in the second half of textbooks;
- The concepts of pronominal verb or pronominalisation are barely explained;
- Focus on formal properties, such as conjugation and subject-verb agreement;
- The different meanings that pronominal verbs can encode are usually not mentioned;
- A couple of textbooks typically refer to daily habits when introducing pronominal verbs.

All in all, there seems to be an overall lack of effort put into the

conceptualisation of this class of verbs, especially when defining the nature of pronominal verbs and the contexts in which they can be used (other than when expressing daily habits). Furthermore, exercises also seem to encourage the memorisation of forms instead of checking whether learners have the ability to decide which form the verb should take. While using descriptive and/or pedagogical grammar treaties to teach pronominal verbs could bring a more precise image of the uses of this concept, they are often neglected by Japanese teachers due to their complexity and the need to adapt the curriculum to the immediate needs of most students (Hirashima, 2010). Indeed, many Japanese universities put the emphasis on the progression in the curriculum to ensure most students are able to graduate after 4 years and join the employment market (Mayaux, 2007); this constraint implies that lessons should be inclusive enough to equally help learners with a low degree of motivation, as well as those who wish to deepen their linguistic knowledge - including their mastery of complex notions such as pronominal verbs. Considering the difficulty of the concept and the methodologies used to teach it, it is no surprise that many learners end up shying away from using pronominal verbs and would rather find substitutes for them, especially when engaged in formal instruction (Nakamura, 2004).

Indeed, the difficulties encountered by Japanese learners in understanding the mechanics behind pronominalisation make it unlikely for them to master this concept well until later stages of linguistic acquisition. As part of a subsidiary study we carried out (Zakhia Doueihi, submitted) while working with a corpus of French learners with 3 different L1, among which Japanese learners, we had the opportunity to extract data regarding the stages of acquisition of French pronominal verbs from a production perspective. Each production contained from 60 to 200 words and were collected in a French

ability test (see Wilkens et al., 2023 for more information about the corpus). The corpus analysis consisted of annotating learners' mistakes of two types when using pronominal verbs: the cases of faulty omissions and faulty additions of the clitic *se* featured with the verbs. We then calculated learners' mean usage rate of pronominal verbs (*PV*) per production and per word. This way, we were able to gather data to apprehend the stages of acquisition with good precision, as each production from learners was labelled with a CEFRL (*Common European Framework of Reference for Languages*) level ranging from A1 to C2. Below is a table summarizing the results of the annotation task for L1 Japanese learners:

CEFRL level	Number of productions	Mean of learners' PV usage rate per production	Mean of learners' PV usage rate adjusted for 100 words production	Number of mistakes
A1	8	0.25	0.39	0
A2	50	0.28	0.36	8
B1	50	0.52	0.57	3
B2	50	0.81	0.67	5
C1	50	1.6	1.07	4
C2	2	1.5	0.94	0

Table 2: Corpus annotation task results for Japanese learners

Although we had limited data in the corpus for levels A1 and C2, a couple observations could still be made through this study. Unsurprisingly, lower to intermediate levels (A2~B2) contained a lower degree of use of pronominalisation, but also more mistakes compared with advanced levels. As learners improve their language skills, they start to be able to conceptualise more L2 structures, and their usage of less formulaic pronominal sentences, in particular, increases considerably. Conversely, it is not surprising that the few uses reported in A1 levels were mistake-free, as novice learners might be prompt to use structures without necessarily

understanding their grammatical functioning, such as *Je m'appelle*.

The results especially demonstrate how actively acquiring pronominal verbs is a long process, as learners seemed to reach a stabilised and balanced usage only from the C1 level onwards. Even though improvements in lexical knowledge – observable in metrics such as CEFRL - are often related to more frequent and adequate use of grammatical content such as pronominal verbs (Miralpeix & Muñoz, 2018), effort put into conceptualisation should alleviate the cognitive process of memorizing each and every pronominal verb from a lexical perspective, so as to provide access to this concept at earlier stages of linguistic acquisition. In particular, we wished to explore whether the involvement of tasks focused on enhancing conceptualisation of pronominal verbs would allow learners to develop their *languageing* abilities and lead to precocious mastery of such a complex, yet key concept. This work will attempt to verify this hypothesis, following the structure detailed in the following last section of this introductory chapter.

1.5. Structure of the thesis

In order to teach pronominal verbs to Japanese university students while allowing them to use these verbs willingly, we advocate that more focus should be put into the understanding of the nature of pronominal verbs and their contexts of use, as their linguistic system does not include a corresponding structure (Nakamura, 2004). In this regard, we believe that the role of their L1 cannot be ignored, as a large part of linguistic knowledge is mediated by our first language (Renoud, 2016). Yet, exclusively focusing on the conceptualisation in L1 and separating the learning process from proper communicative objectives would not give us satisfying results either, as we also think that conceptual knowledge should be linked with practical use.

Taking inspiration from Vygotsky's (1987) socio-cultural theory, we decided to carry out an experiment using Concept-Based Instruction (CBI, also known as *Concept-Based Language Instruction* (CBLI)). In the following chapter, we will present this instruction method, its tenets as well as the influence it had on our research. Especially, we will show that CBI can help in providing tasks focused on conceptualisation of the target concept, which has been lacking. The second part of Chapter 2 will also describe the target concept of French pronominal verbs thoroughly, as well as the inherent difficulties linked to them, while proposing a taxonomy which incorporates most of the pronominal structures which are commonly used.

Following the portrayal of these key theoretical tenets, we will present our research methodology in Chapter 3, which includes the rationale behind the support which was chosen to incorporate CBI along with our comparison treatment, that is to say a digital game. The creation of these games will be detailed upon and motivated by theoretical foundations based on research in Digital-Game (Based) Language Teaching (DGLL). Chapter 4 will then be centered around the description of our practical experimental research which took place in Japanese universities in 2023 and 2024, and which aimed at testing whether a methodology focused on conceptualisation in learners' L1 would bring stronger gains in the acquisition of pronominalisation compared to a task-based approach balanced around language exposure in L2.

Finally, Chapter 5 will cover statistical and qualitative analyses of all gathered results as well as show our interpretations of the findings, before we conclude the thesis in Chapter 6 by offering a summary of the research and question its limits. An overview of the architecture of this work is shown in Figure 5:

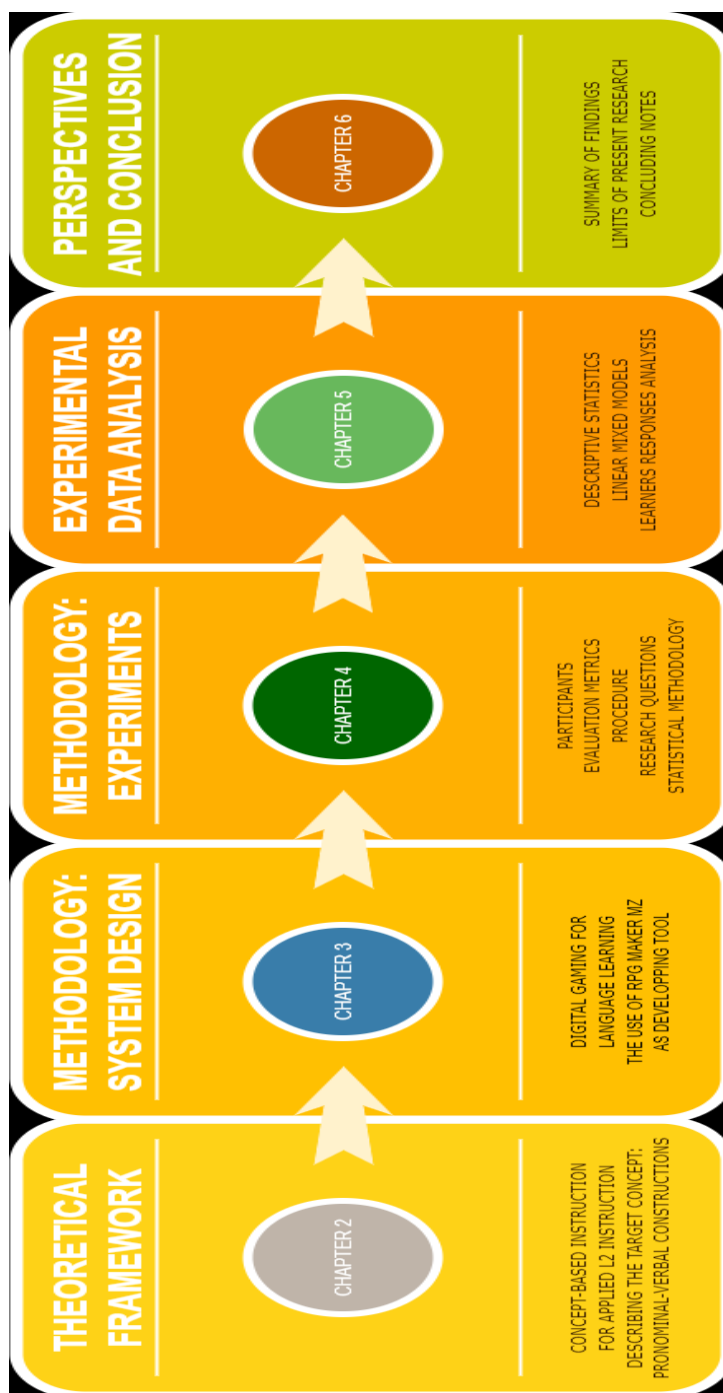


Figure 5: Architecture of this thesis

Chapter 2: Theroretical framework

This chapter will present the theoretical tenets behind our experimental module, notably by offering a systematic review of the reserach in Concept-Based Instruction up to 2025, before describing and proposing a taxonomy of the concept targeted by our experiments, which is to say French pronominal verbs.

2.1. Concept-based Instruction for applied L2 acquisition²

During the last two decades of second language teaching research, concept-based instruction (CBI) has been increasingly promoted as a potentially effective methodology for the development of conceptual skills, especially for grammar-teaching purposes. This methodology gained popularity as L2 instructors realised that "[...] *the quality of explanations provided to L2 learners [...] has a central impact on L2 development*" (Williams et al., 2013, p.365). While there is general consensus on the effectiveness of communicative language teaching methods to quickly enhance second language speaking skills (Richards, 2006), in the case of grammar and the more abstract aspects of linguistic development, the efficiency of such methodologies is more debatable. Direct approaches such as CLT frequently lead to a lack of control over L2 grammatical skills after formal education has ended (Harley et al., 1990; Lyster, 1999, 2004; Swain, 1997). For example, Lyster (2004) noticed that instruction overemphasizing production activities led to difficulties for students in distinguishing the functionality of tenses, especially in the long-term.

Following these criticisms, a second language teaching methodology inspired by Vygotsky's socio-cultural theory (SCT) and called concept-based

² Section 2.1 is based on a publication (Zakhia Doueïhi & François, 2022).

instruction (CBI) has been used in L2 instruction, supporting the use of explicit grammar instruction to enhance control over L2. CBI does not aim to replace CLT but could be seen as complementary to the production-focused methodologies. Essentially, CBI for second language acquisition (SLA) is based on the idea that increased explicit knowledge of difficult concepts (such as grammatical concepts, but also pragmatic, discursive concepts) can not only lead to the development of new cognitive structures, which enable a deeper understanding of the target concepts, but also to long-term acquisition (Lantolf et al., 2015; Ohta & Tsujihara, 2017). It also enhances learners' ability to use the concept, as opposed to the way traditional grammar teaching often explains concepts through incomplete rules of thumb (Garcia, 2017; Negueruela, 2003).

Experimental approaches to CBI are nearly all inspired by Vygotsky's widely known theory of education (1987). They have investigated the learning of different concepts, with the help of distinct methodologies and have returned mixed results. To our knowledge, no systematic review focused on CBI studies for SLA has been made. While we could find reports of previously performed CBI studies in several papers (especially in Kim 2013; Lantolf et al., 2015), there was no specific intent to deliver an exhaustive list for the sake of systematicity. We also wished to produce an analysis including papers reporting both positive and mixed results to give a more faithful overview of the strengths and weaknesses of CBI. This is why we believe that a review of the literature in this promising field of research would not only arouse more interest for researchers in SLA, but would also give further insights into the methodologies used by CBI instructors and the expected benefits of such instruction when used in language instruction.

Our aim to determine the practical utility and degree of applicability of CBI for pronominal verbs among Japanese university students led us to focus our

review on CBI studies which used a classroom intervention. In Section 2.1.1. and 2.1.2., we explore the theoretical background in which CBI is rooted. Section 2.1.3. describes the research methodology that helped us gather CBI studies. Section 2.1.4. is the focal point of the first half of this chapter and offers an overview of the studies analysed according to the criteria described in the research methodology. Finally, in Section 2.1.5, we discuss some remaining challenges and promising avenues for research in the CBI field.

2.1.1. Vygotsky's socio-cultural theory of mind

According to Vygotsky's (1987) socio-cultural theory of mind, development – including learning - is conceptualised as a cognitive process mediated by cultural and social artefacts. In other words, cognitive development is derived from the interactive experience of the subject with the world. This is not to say that human neurobiology has no impact on development, but rather that it is mainly through interactivity that we grow and learn (Lantolf et al., 2015). While engaged in socio-cultural activities (interacting with an item or with other people, for instance), the human brain extends its conceptual layer; humans become able to intentionally control their biological functions - such as the ability to use the hand to hold an item, or the ability to recognise particular symbols (Zhang, 2014). Vygotsky argues that humans grow into the intellectual life of those around them by learning from them, and consequently, that development is essentially a socially constructed process (Lantolf et al., 2015). This is why the help of another person is sometimes necessary – someone more knowledgeable regarding the problem we want to solve. One of the most effective ways of transmitting knowledge is thus achieved in formal education, where one plays the role of an instructor, and the other of a student.

Vygotsky (1987) hypothesises that humans learn most efficiently in the **Zone**

of Proximal Development (or ZPD). The ZPD can be understood as a learning phase situated between the state where a learner cannot solve a problem and the state where he/she can solve it without any help. This phase implies that tasks that learners normally cannot do without help can be achieved with enough assistance from an instructor. Connected to the importance given by Vygotsky to social interaction in the developmental process, the concept of ZPD asserts that while we can learn to some extent by experiencing the world's objects ourselves, it is less likely that mathematical concepts, for instance, can be learned without the aid of an instructor (Crain, 2010). To acquire a new complex process, students should be helped in a way that their current knowledge allows them to understand most of what is being taught, but still requires a further step in order to master the new concept – granted by the instructor's teaching. The ZPD, once crossed, then leads to a newly acquired developmental layer where the student is independent. Development becomes a progressive process where learners are continuously approaching their next abilities as they try to solve new problems (Lantolf et al., 2015).

Since we learn most complex concepts through the help of an instructor in an educational setting while using constructed forms of mediation to interact with the world (language, for instance), we also gain access to what Vygotsky calls “**scientific concepts**”. Scientific concepts (as opposed to everyday concepts) are general concepts that are learned systematically and not through empiricism. As we begin to internalise these concepts, we become able to transfer their principles to any necessary context. The goal of school instruction is thus to gain control over those scientific concepts (Negueruela & Lantolf, 2006).

2.1.2. Concept-based instruction

The importance given to the knowledge of scientific concepts by SCT is one of the main tenets of concept-based instruction for second language teaching. According to CBI, most language teaching methods rely on the use of rules of thumb (such as: “Spanish preterit should be used when an action was completed once in the past”) to teach grammatical notions, but these rules lack the systematicity of scientific concepts (Garcia, 2017). This can lead to incorrect uses of the notions and overgeneralisation and, as Pavlenko (1999) claims, while students can provide the definition of a grammatical concept, it does not mean they have control over its functionality. After asking his students to write arbitrary properties of two Spanish past tenses, Negueruela (2003) reported that in some cases, their definitions did not even match the examples they used to illustrate these properties. This shows that there might be a lack of control in how grammatical concepts are learned and perceived in traditional grammar instruction.

In a CBI classroom, the focus is shifted from memorizing rules to understanding the underlying concepts of an abstract grammatical notion. While this requires an exhaustive understanding of the grammatical notion and a dedicated teaching curriculum from the teacher, it also enables learners to mediate their own cognitive activity, and thus independently apply their new understanding to larger situations (Lantolf, 2011).

To promote the teaching of scientific concepts in classroom instruction, Gal’perin (1989, 1992) devised a methodology called STI (“Systemic-theoretical instruction”). This model, one of the first CBI models applied to second language teaching and, more generally, formal instruction, follows a three-step procedure: **materialisation**, **verbalisation**, and **internalisation**.

The **materialisation** phase first aims to create a representation of the taught concept. Concretely, a visual representation such as a diagram or a graph is

used to depict the concept in an exhaustive and systematic way, because Gal'perin believes an image has more impact on the students' representations and is more easily grasped than a verbal definition. During this phase, the teacher must explicitly give access to the concept by explaining the constructed scheme, or SCOBAs (*Scheme of a Complete Orienting Basis of the Action*). By understanding the scheme, learners should have a better idea of how the concept works, and it should be designed so that any choice the student makes in the use of the grammatical concept is rationally motivated (Gal'perin, 1989). While the form of this SCOBAs can vary (ranging from flowcharts to 3D objects) it is extremely important for the taught concept to be as comprehensively presented as necessary. This step could be considered as a first scaffolding in the concept acquisition process that should eventually allow unmastered knowledge to be assimilated.

The **verbalisation** phase then requires learners to explain the concept. This not only helps learners to verify their own understanding of the concept, but the activity itself is also a way of enhancing their ability to use the concept. From the teacher's perspective, this phase involves the monitoring of the verbalisations, whether they be made in a classroom or through homework recordings. Bringing concepts to consciousness by "linguaging" allows students to recontextualise the new conceptual knowledge to other contexts and enhance their overall control over the notion (Lee, 2016).

Finally, the **internalisation** phase is considered the objective of the whole sequence. Once the concept is internalised, it means learners are able to fully comprehend and apply the concept to any context without guidance of the SCOBAs – and then, overcoming the Zone of Proximal Development. Measuring conceptualisation and/or internalisation is essential because it encapsulates how scientific concepts, also known as declarative knowledge in cognitive linguistics, developed into procedural knowledge – akin to the

practical, strategic use of the target concept, a factor usually underevaluated in Vygotsky's social-cultural theory (Karpov, 2003; Star, 2005).

While a few attempts to implement Gal'perin's theory in L2 instruction have been carried out (Carpay, 1974; Karpova, 1977), it was Negueruela (2003) who first extensively addressed STI to perform a CBI classroom study in everyday instruction. In order to teach Spanish verbal mood, tense, and aspect to L1 English/L2 Spanish learners, the researcher implemented interactive instruction based on in-class discussions, at-home audio-recorded verbalisations and the use of a flowchart, which worked as the SCOA to explain the concepts. The aim of this study was to allow students to use these grammatical concepts according to their communicative intents rather than based on incomplete rules of thumb. The teaching methodology widely followed STI's main principles, with a strong focus on the importance of verbalisation activities (6 at-home audio-recorded tasks were carried out) to internalise the studied concepts. The study demonstrated how students were able to make *“coherent connections between accurate performance using preterit and imperfect and conceptualisation of the meaning of preterit and imperfect through the notion of verbal aspect”* (Negueruela, 2013, p.235).

Following the success of Negueruela's work, several attempts to apply a similar concept-based grammar teaching methodology were carried out. These studies progressively made up the field of CBI for SLA. The growing body of research in this field makes it timely to give systematic reviews of the domain to facilitate experiments based on common tools. This is the reason why we aimed, in the present chapter, to gather and review all studies that intended to use a concept-based method for the acquisition of grammatical concepts within a classroom setting.

2.1.3. Research methodology

In order to provide the most exhaustive possible survey of studies about experimental Concept-based Instruction in SLA, we applied the principles used in scoping reviews (Kitchenham & Charters, 2007). These principles include three steps: (1) identifying relevant literature using a systematic search strategy; (2) selecting the corpus of studies based on inclusion and exclusion criteria and (3) extracting and coding variables for further analysis and interpretation.

In the first step, we carried out a systematic search based on common English keywords in the field of CBI for SLA in 5 meta-databases. The databases chosen were CAIRN, De Gruyter, ProQuest, MLA and ScienceDirect, and were all chosen for their relevance in the field of second language teaching. The keywords that we considered likely to address most of the studies we were looking for are: Vygotsky / SCT (*Socio-cultural theory*) / STI (*Systemic-theoretical Instruction*) / CBI / “Concept-based instruction” / Concept / L2 / Second Language / FLE (*Français Langue Étrangère*) / ESL (*English as Second Language*). We then performed several searches using combinations of two, three, or four of these keywords in the aforementioned databases via advanced searches. We obtained varied results, ranging from 6 studies to 3,502 studies. A manual investigation of the distributions of these results allowed us to pinpoint a set of best keywords to get the maximal number of hits related to SLA. The results for the five best sets obtained at the end of this manual analysis are displayed in Table 3.

Keywords combination	Database	CAIRN	De Gruyter	ProQuest	MLA	Science Direct	Total number of hits
Vygotsky AND Concept AND (Second Language OR L2) AND CBI		0	22	31	6	14	73
Vygotsky AND Concept AND (Second Language OR L2) AND STI		0	5	25	1	16	47
Vygotsky AND Concept AND (Second Language OR L2) AND ESL		5	292	605	6	220	1128
Vygotsky AND Concept AND (Second Language OR L2) AND SCT		0	68	57	15	51	191
"Concept-based instruction" AND Concept AND (Second Language OR L2)		0	66	32	20	18	136

Table 3: Number of hits per keyword combination (restricted search)

While the combination of the keywords Vygotsky / Concept / Second Language / ESL returned the most hits, it was still not restrictive enough to faithfully cover the field of SLA CBI studies and also had the disadvantage of omitting studies performed with other languages as L2. We thus decided to focus on the second and third largest set of keywords, which were the combination of keywords Vygotsky / Concept / Second Language OR L2 / SCT (191 papers), and the combination of keywords "Concept-based instruction" / Concept / Second Language OR L2 (136 papers).

As a second step, we analysed all the papers found to dismiss those which did not fall within the scope of our survey. In accordance with scoping review methodology, we relied on a set of inclusion and exclusion criteria. Inclusion criteria were that studies had to be about CBI. Secondly, CBI had to be applied to SLA. Finally, only studies conducted after 2003 were reviewed, as the field of CBI for SLA started to gain interest mostly after Negueruela's (2003) study. As regards exclusion criteria, we excluded articles written in languages we could not understand. As such, studies in English and French were included, while others were excluded from our

survey. Secondly, we excluded research which did not implement a classroom experiment, to better assess the effects and limitations proper to CBI. Application of the aforementioned criteria resulted in a corpus of 20 studies, distributed among the 5 databases as in Table 4:

Studies	CAIRN	De Gruyter	ProQuest	MLA	ScienceDirect
Negueruela (2003)	Ø	Ø	Ø	√	Ø
Twyman et al. (2003)	Ø	Ø	√	Ø	Ø
Swain et al. (2009)	Ø	√	Ø	Ø	Ø
Gánem-Gutiérrez & Harun (2011)	Ø	√	√	√	Ø
Garcia (2012)	Ø	√	Ø	√	Ø
Lee (2012)	Ø	Ø	Ø	√	Ø
Van Compernelle (2012)	Ø	√	Ø	√	√
White (2012)	Ø	√	Ø	Ø	Ø
Kim (2013)	Ø	√	Ø	√	Ø
Polizzi (2013)	Ø	Ø	Ø	√	Ø
Harun et al. (2014)	Ø	Ø	Ø	Ø	√
Zhang (2014)	Ø	Ø	Ø	√	Ø
Aguilo-Mora & Negueruela(2015)	Ø	√	Ø	Ø	Ø
Buescher & Strauss (2015)	Ø	√	Ø	Ø	Ø
Masuda & Labarca (2015)	Ø	√	Ø	Ø	Ø
Fehrmann (2016)	Ø	√	√	Ø	Ø
Fogal (2016)	Ø	Ø	Ø	√	Ø
Williams (2016)	Ø	√	Ø	√	Ø
Ohta & Tsujihara (2017)	Ø	√	Ø	Ø	Ø
Suñer & Roche (2019)	Ø	√	Ø	Ø	Ø

Table 4: Studies distribution within the 5 databases

This search in the databases (originally carried out in 2019 as part of a publication) allowed us to gather a solid corpus, but we went through two additional search steps. The first step consisted of a manual investigation of the citation network of the 20 selected papers. This process was twofold: first, we used a forward citation search via Google Scholar to find more recent studies; then we carried out an ancestry citation search to find older articles. Finally, we manually searched for notable studies published between 2019 and 2025. This led to the inclusion of fourteen additional studies, yielding a total of 34 studies published between 2003 and 2025. These studies were analysed according to the following coding scheme.

We identified three main dimensions to code: the context; the material analysed; and the experimental design. The exact scheme used to code for these dimensions is the following. The first dimension is concerned with the context of the study and covers the research duration and the study sample. The duration refers to the time interval between pre-test(s) and post-test(s) as well as, when mentioned in the studies, the frequency of the classroom sessions. The sample will also be described in terms of sample size and, for each study, the L1/L2 of the participants (in some cases, there are several L1s in the classroom).

The first-dimension variables and values are:

Duration: categorical (nominal) and numeric variable (course unit, day, week, semester, year).

Number of participants: numeric variable.

Participant L1: categorical (nominal) variable.

Participant L2: categorical (nominal) variable.

The second dimension corresponds to the material analysed by the researchers/instructors, and includes the concepts taught as well as developmental data. Developmental data includes output variables that are closely linked to the instruction module, which is why most of them are in-class tasks. We used generic terms to sum up the wide variety of activities that were performed in class, while making distinctions between tasks of different nature (e.g. definitions vs compositions). This dimension gives information about the methodology used by the researchers/instructors in their attempt to use CBI in a classroom experiment.

Second dimension variables and values are:

Aimed concept: categorical (nominal) variable.

Developmental data: categorical (nominal) variable (i.e. definition,

composition, written task, verbalisation)

The third dimension describes the nature of the examination of the CBI module performed by the researcher. It includes the way the experimental design was established (within-subject design or between-subject design), whether statistical testing was performed (if no statistical testing was used, the alternative evaluation process will be mentioned) and what were the results. This last variable also codes whether a learning effect was observed at the end of the experiment, and whether that improvement was significant or not from a statistical point of view.

Third dimension variables and values are:

Design: categorical (nominal) variable (Repeated Manner (RM), Independent Groups Repeated Manner (IGRM)).

Statistical testing: y/n + categorical (nominal) variable (in case of absence of statistics).

Improvement: categorical (nominal) variable, including:

- **No improvement**: pre- and post-test values compared without using a statistical test and reported no improvement;
- **No effect**: statistical testing conducted, but no significant effect detected;
- **Improvement**: pre- and post-test values compared without using a statistical test and reported improvement;
- **Significant improvement**: statistical testing yielded a significant effect.

2.1.4. Overview of studies carried out from 2003 to 2025

In this section, we present the results of our systematic analysis based on a corpus of 34 CBI studies, published between 2003 (marked by Negueruela's seminal study) and 2025 (as the year we defended this thesis). The analysis is based on the coding of various characteristics of these studies, grouped according to the aforementioned three dimensions. For each dimension, a table containing the coded characteristics of all studies is reported (the studies are presented by chronological order) and we discuss the most important elements emerging from our analysis.

Research	Context			
	<i>Duration (frequency)</i>	<i>Participants</i>	<i>L1</i>	<i>L2</i>
NEGUERUELA (2003)	16 weeks	12	English	Spanish
TWYMAN ET AL. (2003)	1 course unit	34	Spanish	English
FERREIRA (2005)	16 weeks	14	7 diff. languages	English
YANEZ-PRIETO (2008)	1 semester	14	English	Spanish
THORNE ET AL. (2008)	1 semester	118	Several diff. langu.	English
SERRANO-LOPEZ & POEHNER (2008)	3 weeks	80	English	Spanish
SWAIN ET AL. (2009)	1 week	10	English	French
GÁNEM-GUTIÉRREZ & HARUN (2011)	2 days	6	Arabic, Thai	English
VAN COMPERNOLLE (2012)	6 weeks	8	English	French
GARCIA (2012)	12 weeks	32	English	Spanish
LEE (2012)	6 weeks (1x/week)	23	Chinese, Korean, Thai	English
LAI (2012)	8 weeks (5x/week)	16	English	Chinese
WHITE (2012)	7 weeks	30	7 diff. languages	English
KIM (2013)	12 weeks	9	Korean	English
FRAZIER (2013)	1 semester (2x/week)	6	English	Spanish
POLIZZI (2013)	1 year	1	Ukra., Russ., English	Spanish

HARUN ET AL. (2014)	1 course unit	8	Malay	English
ORD (2014)	7 weeks	5	English	French
ZHANG (2014)	2 weeks	6	English	Chinese
BUESCHER & STRAUSS (2015)	2 days – 3h total	8	English	French
AGUILO-MORA & NEGUERUELA (2015)	1 course unit	20 (14 + 6) ³	English/Spanish	Spanish/English
MASUDA & LABARCA (2015)	1 week	8	English	Japanese
RENOUD (2016)	21 weeks - first exp	7	Japanese	French
RENOUD (2016)	12 weeks - sec. exp	5	Japanese	French
WILLIAMS (2016)	1 day	12	English	French
FEHRMANN (2016)	4 weeks	10	Japanese	German
FOGAL (2016)	8 weeks – 16h total	7	Japanese	English
FAZILATFAR ET AL. (2017)	8 weeks (1x/week)	28	Iranian	English
OHTA & TSUJIHARA (2017)	8 weeks	12	English	Japanese
SUÑER & ROCHE (2019)	One 1h30 session	39	French	German
LIU & HSIEH (2020)	4 weeks – 4h total	3	English	Chinese
BENHAMLAOUI & GÁNEM-GUTIÉRREZ (2022)	3 weeks – 9h total	45	Arabic	English
HESS & AMORY (2022)	6 days – 5h total	7	English	Korean
QIN ET AL. (2023)	5 weeks	53	Chinese	English
WANG (2023)	40 weeks	8	Chinese	English

Table 5: Study context

CBI experiments conducted between 2003 and 2025 in second language classroom environments range from a single session (intensive classroom lesson carried out by Suñer & Roche, 2019) to a whole year (developmental case study conducted by Polizzi, 2013). There seems to be no common ground on whether a CBI module should be carried out in the short-term or

³ Aguilo-Mora & Negueruela (2015) performed an experiment in 2 classrooms in parallel, one with 14 students learning Spanish as L2 and the other one with 6 students learning English as L2.

in long-term, although it is important to note that many researchers stated a longitudinal study would be beneficial after conducting their experiment, even when results were positive (Buescher & Strauss, 2015; Frazier, 2013; Lai, 2012; Ord, 2014; Williams, 2016). In some studies (Aguilo-Mora & Negueruela, 2015; Harun et al., 2014; Twyman et al., 2003), the study duration was not mentioned.

The frequency of classrooms session is also a factor that must be considered. While most of the research took place instead of the regular language lesson, some experiments were intensive and extended the duration of the curricular lesson (Lai, 2012; Williams, 2016). It is yet to be determined whether classes should be given extensively over a short period of time, or if there are merits in leaving enough time between each lesson for students' memories to internalise.

The number of participants is also highly variable, ranging from a single student (Polizzi, 2013) to 118 students (Thorne et al., 2008). Out of the 34 studies reviewed, 8 were carried out with over 30 students, and 17 were conducted with 10 students or less. These numbers are in line with the general mean of students in between-subject designs in classroom experiments of 24.6 (Lindstromberg, 2016).

A review of participants' target language shows that among the 34 studies reviewed, 14 were conducted with students learning English as an L2, 7 with Spanish, 6 with French, and 3 with Chinese. As regards student's L1, English dominated (L1 of participants in 18 of 34), most likely due to the fact that most of these studies were carried out in North America. East-Asian languages (Chinese, Japanese and Korean) are also moderately represented as the L1 of learners in 8 studies. Spanish and Thai are each represented in 2

studies. It is worth mentioning that 6 studies were performed in classes with mixed L1 backgrounds, in which case, the teaching intervention was entirely carried out in the learners' L2.

Research	Material	
	<i>Aimed concept</i>	<i>Developmental data</i>
NEGUERUELA (2003)	Aspect, mood, time	Definitions, in-class performances, individual recordings
TWYMAN ET AL. (2003)	Civilisations	Compositions, quiz, MCQ, map creations, essays
FERREIRA (2005)	Writing genres	In-class performances, in-class interactions, problem-solving tasks
YANEZ-PRIETO (2008)	Aspect, time, textual properties	In-class performances, compositions (3), interviews (2)
THORNE ET AL. (2008)	Deontics, directive speech	Roleplay, performances, interviews, corpuses
SERRANO-LOPEZ & POEHNER (2008)	Locative prepositions: <i>de, en, a</i>	Written tasks, verbalisations
SWAIN ET AL. (2009)	Voice	Definitions, writing tasks, individual recordings
GÁÑEM-GUTIÉRREZ & HARUN (2011)	Aspect	Definitions, written tasks (exercises), oral verbalisations (in pairs)
VAN COMPERNOLLE (2012)	Sociopragmatics	Interviews (6), questionnaires, scenarios
GARCIA (2012)	Aspect	Definitions, in-class performances, individual recordings
LEE (2012)	Phrasal verbs: <i>out, up, over</i>	Definitions, compositions, verbalisations, MCQ, interviews
LAI (2012)	Temporal system, aspect markers	Definitions, in-class performances, compositions, dialogues, translations
WHITE (2012)	Phrasal verbs: <i>up, out, through, off, down, in</i>	Definitions, written tasks (explaining meaning), drawings, verbalisations (disc.)
KIM (2013)	Sarcasm	Interviews, in-class performances, compositions
FRAZIER (2013)	Verbal mood	Definitions, written and oral tasks (compositions, recordings)
POLIZZI (2013)	Aspect	In-class performances, verbalisations, interviews
HARUN ET AL. (2014)	Aspect, time	Verbalisations (language units)
ORD (2014)	Aspect	Interviews (3), written and oral tasks, discussions
ZHANG (2014)	Topicalisation	Definitions, compositions, in-class performances
BUESCHER & SRAUSS (2015)	Locative prepositions: <i>à, dans, en</i>	Written tasks, compositions, discussions/verbalisations
AGUILO-MORA & NEGUERUELA (2015)	Motion verbs	Written tasks (translations), oral tasks (giving directions), verbalisations
MASUDA & LABARCA (2015)	Particles <i>ni, de</i>	Written tasks (narratives), definitions, verbalisations
RENOUD (2016)	Aspect, time	Definitions, written tasks, compositions, individual recordings
WILLIAMS (2016)	Auxiliary verbs	Written tasks (auxiliaries choice), verbalisations (Chatzy platform)
FEHRMANN (2016)	Declarative sentences: prefield (V2 structure)	Written tasks (writing mails), verbalisations, in-class performances

FOGAL (2016)	Academic writing: stylistics, voice	Written tasks (essays, stylistic worksheets), recorded discussions, interviews
FAZILATFAR ET AL. (2017)	Aspect, time	Definitions, compositions, multiple-choice questions, verbalisations (written)
OHTA & TSUJIHARA (2017)	Honorific speech (<i>wakimae</i>)	Roleplay, written tasks, individual recordings, online surveys, discussions
SUNER & ROCHE (2019)	Light-verb constructions	Written tasks (fill-the-gaps, reformulations, cover letter)
LIU & HSIEH (2020)	Chinese metaphor	Written tasks (fill-the-gaps, word cards), video watching, drawings, narrative creating, verbalisations
BENHAMLAOUI & GÁNEM-GUTIÉRREZ (2022)	Aspect, time	Practice tasks, recorded discussions, verbalisations (Dialogic thinking)
HESS & AMORY (2022)	Honorific speech (<i>hearer honorifics, speech ending style and subject honorifics</i>)	Written tasks, recorded discussions, role-play, skit script, recorded verbalisations
QIN ET AL. (2023)	Modal verbs	Written tasks, verbalisations (written), oral languaging tasks (SCOBAs)
WANG (2023)	Conceptual metaphor	Written tasks (Google Docs), recorded discussions, journal entries, video recordings

Table 6: Study material

The analysis of the concepts that were targeted by each study shows a clear trend towards the teaching of aspectual properties: out of the 34 studies reviewed, 11 studies focused on aspect. This might come from the fact that Negueruela's seminal study aimed to teach Spanish aspect. Aspect is also widely acknowledged as a difficult grammatical concept to acquire, especially in a teaching context where there are inadequacies in the way grammatical explanations of aspect are presented in textbooks (Negueruela & Lantolf, 2006).

Another interesting category of concepts that was researched are tied to spatial properties, such as locative prepositions (2 studies) and, to some extent, phrasal verbs (2 studies). In that regard, Serrano-Lopez & Poehner's (2008) attempt to allow learners to internalise locative prepositions referring to spatial concepts was especially innovative, as the researcher turned her teaching module into a practice session with 3D clay models.

Besides grammatical concepts, some researchers applied a concept-based method to concepts linked to other linguistic categories, such as pragmatics (distinction between French forms of address by Van Compernelle, 2012; sarcasm by Kim, 2013; topicalisation by Zhang, 2014; honorific speech by Ohta & Tsujihara, 2017 and by Hess & Amory, 2022). There is also a case where a purely semantic concept was studied (the concept of civilisations, see Twyman et al., 2003). There seems to be no particular constraint with carrying out a CBI with concepts that are not purely grammatical.

In order to gather developmental data from their experiments, researchers mostly adopted the three following teaching strategies. The most common way was a definition task, in which students must explain the concept. Most researchers resorted to this task either directly during interviews or through a written task. Another commonly used process was the recordings of students' verbalisation. Finally, one of the most indicative markers of improvements in concept acquisition consisted in performance data obtained through the use of in-class direct observation.

Research	Experimental design		
	<i>Design</i>	<i>Statistical testing</i>	<i>Results</i>
NEGUERUELA (2003)	RM	No (qualitative analysis of definitions and discourse)	Improvement
TWYMAN ET AL. (2003)	IGRM	Yes	Significant improvement
FERREIRA (2005)	RM	Yes	Significant improvement
YANEZ-PRIETO (2008)	RM	No (test score contrastive analysis)	Improvement
THORNE ET AL. (2008)	RM	No (correct usage rate comparison)	Improvement
SERRANO-LOPEZ & POEHNER (2008)	IGRM	No (test score contrastive analysis)	Improvement
SWAIN ET AL. (2009)	RM	Yes	Significant improvement
GANEM-GUTIÉRREZ & HARUN (2011)	RM	No (test score contrastive analysis)	Improvement
VAN COMPERNOLLE (2012)	RM	Yes	Significant improvement
GARCIA (2012)	RM	No (qualitative analysis of definitions)	Improvement
LEE (2012)	RM	Yes	Significant improvement
LAI (2012)	IGRM	Yes	Significant improvement

WHITE (2012)	RM	Yes	No effect
KIM (2013)	RM	Yes	Significant improvement
FRAZIER (2013)	RM	No (test score contrastive analysis)	Improvement
POLIZZI (2013)	RM	Yes	Significant improvement
HARUN ET AL. (2014)	RM	Yes	Significant improvement
ORD (2014)	RM	Yes	Significant improvement
ZHANG (2014)	RM	Yes	Significant improvement
BUESCHER & STRAUSS (2015)	RM	Yes	Significant improvement
AGUILO-MORA & NEGUERUELA (2015)	RM	No (qualitative analysis of the correctness of translations)	No improvement
MASUDA & LABARCA (2015)	RM	Yes	Significant improvement
RENOUD (2016)	RM	Yes	Significant improvement
WILLIAMS (2016)	RM	Yes	Significant improvement
FEHRMANN (2016)	IGRM	Yes	Significant improvement
FOGAL (2016)	IGRM	Yes	No effect
FAZILATFAR ET AL. (2017)	IGRM	Yes	Significant improvement
OHTA & TSUJIHARA (2017)	RM	No (qualitative analysis of the correct understanding of the concept)	Improvement
SUNER & ROCHE (2019)	IGRM	Yes	Significant improvement
LIU & HSIEH (2020)	RM	No (qualitative analysis of the correct understanding of the concept)	Improvement
BENHAMLAOUI & GÁNEM-GUTIÉRREZ (2022)	IGRM	Yes	Significant improvement
HESS & AMORY (2022)	RM	No (test score contrastive analysis + qualitative analysis of the correct understanding of the concept)	Improvement
QIN ET AL. (2023)	IGRM	No (qualitative analysis of the correct understanding of the concept)	Improvement
WANG (2023)	RM	No (qualitative analysis of the correct understanding of the concept)	Improvement

Table 7: Experimental design

The experimental design of CBI studies is also a key characteristic as it is a clue about the confidence one can have in the results reported in the study. Two types of experimental designs were used: repeated manner (RM) and independent groups with repeated manner (IGRM). RM designs measure the effects of the training at different times within the same group, while IGRM

designs not only measure effects of CBI training as the difference between post-experimental performance and pre-experimental performance, but they also compare this performance with that of a comparison group. We found 9 studies carrying out an IGRM design, the other 25 had no comparison group. The 9 studies using IGRM design compared the CBI group with a comparison group taught following a traditional approach. Among them, 3 studies stand out as they performed a multidimensional design (having more than one comparison group as a basis of comparison) that could be inspiring for future research in the field. Serrano-Lopez & Poehner (2008) investigated the impact of the nature of the CBI modules through an experiment with three groups: a CBI module including the use of 3D clay models as a visual tool, a CBI module without the use of clay, and a traditional communicative approach module. Lai (2012) compared the effects of a CBI module on elementary-level students with two comparison groups receiving traditional teaching: one of elementary-level students and one of intermediate students. This design helped to better characterise the conceptual development in the experimental group compared to similar and more advanced learners. Going further on this track, Fehrmann (2016) tested the effects of a CBI module for German L2 in comparison with a comparison group of learners receiving no CBI, but also with a comparison group composed of native German speakers, representing the mastery native speakers have of the concept.

Twenty-one studies carried out statistical testing. In cases where no statistical testing was performed, the researchers used various methods to analyse the data, resorting to using purely qualitative analysis (7 studies) or comparing test scores of the samples without using a statistical test to generalise (6 studies).

The last issue we address is the effectiveness of the CBI approach in terms of conceptual development. Most studies (31 out of 34) show, at various degrees, overall improvements in the conceptual knowledge and understanding of the target concept. It is interesting to note that the majority of the studies where results were statistically tested showed significant improvements in the conceptual development, with only 2 studies reporting no effect of CBI. Calculations of Cohen's *d* effect sizes demonstrate significant effects within CBI groups, with magnitudes from pre- and post-tests comparison as high as 0.58 (Fazilatfar et al., 2017), 0.67 (Lee, 2012), 1.12 (Suñer & Roche, 2019) or even 2.15⁴ (Kim, 2013). Studies where developmental data were analysed qualitatively and/or without statistical testing also reported improvement, apart from Aguilo-Mora & Negueruela (2015). Among the two other studies reporting no effect, White (2012) notes that the results he obtained with this method were not significant, pointing out the fact that the phrasal verbs items which were asked in the tests were not treated in the intervention. The task at hand may then have been too difficult at that point in time. Likewise, Fogal (2016) reports no differences between groups and slight to moderate effects in the long term. The reason advanced by the researcher was that the target concept (academic writing proficiency) requires years to develop, so that the three-week intervention was insufficient for internalising the target discourse features. Also, it should be noted that the results reported in Fogal's (2016) study only involved two groups of respectively 7 (for the intervention group) and 3 (for the control group) learners, which could hinder the statistical reliability of these results.

It is noteworthy that Serrano-Lopez & Poehner (2008), Lai (2012), and Fehrmann (2016) all reported positive results from their multidimensional

4 This could be explained by the fact that the difference in results found in pre-test were important and by the low number of participants ($n = 9$)

studies. Serrano-Lopez & Poehner's module showed that the CBI+clay condition outperformed the two other groups, especially in the long term. In Lai's study, the experimental group not only performed better as regards concept mastery than the comparison group with the same initial level, but also performed no differently than the more proficient intermediate level comparison group. Fehrmann (2016) also reports positive effects of the CBI module, which allowed learners to reach levels of understanding of the concept close to that exhibited by L1 German speakers. The latter experiments suggest that CBI instruction could have a stronger effect than other teaching methods on the conceptual knowledge of learners.

It is important to note that although results were generally positive, researchers should still consider these results with caution as experiments with CBI can lead to mixed results (as seen in 3 studies reviewed in this chapter). Furthermore, the field could also be subject to **publication bias**. As such, publications with negative results could have been more difficult to identify or might not have been published at all.

2.1.5. Challenges and avenues for future research in CBI

The overview of the CBI field of research sheds light on the current state of development of the domain. Ever since the core study performed by Negueruela (2003), STI-inspired concept-based pedagogy has grown in popularity, and has essentially focused on an alternative way to teach complex grammatical concepts whose acquisition is often hampered by traditional instruction manuals that generally explain grammatical concepts using rules-of-thumb.

Although most previous studies generally bear witness to the success of CBI

in SLA grammar acquisition, it is still hard to say whether the method itself could become standard in L2 curricula. Current limitations of CBI should be balanced against the potential benefits granted by a conceptual approach in L2 acquisition. A major concern comes from the fact that only a minority of the reviewed studies (around 25%) included a comparison group. It is therefore hard to judge whether CBI is more effective than a traditional method. Additionally, several researchers argue that longitudinal studies should be more frequent as they would allow deeper research and enable the observation of more data (Buescher & Strauss, 2015; Frazier, 2013; Masuda & Labarca, 2015; Ord, 2014; Swain et al., 2009; Williams, 2016). Third, the limited sample sizes used in most studies raise concerns about the generalizability of the results. Even when statistical tests were performed, experiments were often carried out in classes with a number of students ranging between 10 and 30, which could be considered insufficient for statistical reliability. A solution could be to carry out several replication studies and combine their effect sizes through a meta-analysis, to get a more reliable estimator of CBI effect size.

Another concern comes from the fact that creating SCOBAs to materialise complex grammatical knowledge is not an easy task. As Fogal (2016, p.180) noted after his experiments, CBI *"raises questions about the effectiveness of [SC]OBAs—either because they may be poorly-conceived or because learners do not perceive their (intended) effect — for concept development."* Additional guidance for the instructors seems to be required, as, ideally, the knowledge of the target concept should be deep enough to construct relevant SCOBAs. While in principle, any concept could be materialised, it is worth noting that the more complex a concept is, the heavier the cognitive load required for students and the deeper the degree of involvement with the

target concept will be. Notably, the lack of reported studies applied to the development of traditional communicative abilities, such as listening, speech or oral fluency speaks to the fact that some concepts might prove to be difficult to represent with a SCOBAs without inventivity and adaptability (Tang & Wang, 2024).

Given that CBI tries to explain the concepts thoroughly, it also involves a huge investment of time and energy for teachers and students to learn a specific concept, which can potentially lead to a slowdown in curriculum advancement. Reported studies such as Fogal's (2016) work on the effectiveness of CBI on the acquisition of academic writing showed that even with a 16 total hour of intervention, the concept was not mastered completely and learners did not perceive as though they improved much. In turn, allocating such an important duration of the curriculum for CBI practices on a single concept could/should not be made in any given context. CBI should most and foremost be considered for learners wishing to have a higher degree of awareness, especially regarding grammatical notions (Kambaru, 2023). As such, concept-based instruction agenda should carefully be designed beforehand, so that the practice of communicative skills is not entirely replaced by time-consuming CBI activities.

Bearing these limitations in mind, several axes of improvement were investigated for our research. First, in order to assert the use of CBI, the solid SCT/STI theoretical basis that sustains it should be supported by instruction manuals and curricula, provided these are based on SCOBAs. While SCOBAs are by nature adapted to each concept they should describe, their construction is often similar (a graph or diagram of all iterations of a given concept). A more thorough, explicit treatment would probably benefit

learners who would want to dive deeper into their second language, and we believe that this is something SCOBAs can provide. However, in order to create such agendas, SCOBAs must have been tested and their systematic efficiency should have been demonstrated. This requires an emphasis on reliability in the experimental stages, including pilot studies and power analyses based on the target population.

Another key to the potential success of CBI would come from the use of new technologies. The work of Serrano-Lopez & Poehner (2008) took advantage of technology to provide 3D modelling using clay as a physical support, while Suñer & Roche (2019) made use of multimedia animations to show the uses of light-verb constructions⁵. In both cases, the use of these new supports proved their usefulness as the results were, in the long-term, better for the experimental group than for the comparison group. Further uses of digital technologies, in particular virtual environments, could yield more interactive SCOBAs, and also be easier to share. Allowing learners to suggest their own conceptual mapping through a MOOC or an app could not only lead them to train their understanding as part of verbalisation, but also allow other learners to react and offer their own vision of the concept in the process. Additionally, digital data also has the advantage of being easier to gather, allowing for quicker collection of results.

Strengthening the links between CBI and the field of Cognitive Linguistics (CL) might also have positive impacts on the domain. As Robinson & Ellis (2008, p.7) asserted, "[...] *some aspects of an L2 require awareness and/or attention to language form – implicit learning is not sufficient for successful*

5 Light-verbs are, in German, verbs that do not have a strong intrinsic meaning and work together with another expression, e.g. *nehmen* in *Angriff nehmen* (lit. *to take attack*, meaning “to attack”).

SLA [...]". In CL theories, linguistic knowledge is seen as a part of global cognition and can be learned through the development of categories. Categories can be seen as networks of concepts that humans create progressively and that allow us to relate to more and more difficult meaning-demanding tasks. Consequently, meaning is seen as the key to linguistic development, and explicit language teaching is not seen as redundant, a view shared by the SCT and CBI. In that regard, Lantolf (2011) argues that CL can notably play a role in the initial phase of CBI, the materialisation phase: just as categories in CL are materialised through visual representations and schemes, concepts in CBI take the form of SCOBAs. Using theories of categorisation might prove to be useful in the creation of adequate SCOBAs. This includes making use of the L1 as the main way of *linguaging* – and thus, for the SCOBAs, because the creation of new categories is usually made in the tongue learners are most comfortable with (Kambaru, 2023).

CL also argues that cognition is embodied. This means that concepts, categories, reasoning, and judgment are all connected to certain parts of the body - mostly the neural structure (Borghi & Cimatti, 2010). This view is shared by neuroscientists who have shown that sensory structures of the brain are connected to more abstract mental functions such as consciousness, emotion, self-awareness, and will (Damasio, 1999; Edelman, 2004). While SCT is mostly focused on the social component of development, it does not exclude the impact of neurobiology, and uncovering the effects of embodiment and how conceptualisation is performed in the brain might be useful to create effective SCOBAs. It could be possible to deliver more efficient forms of conceptualisation tools by taking these assets into account (as we could see, for instance, that visual representations often had better long-term effects and deeper understanding in CBI experiments).

2.1.6. Conclusion

After taking all those elements into consideration and carrying out this systematic review of literature, CBI appealed to our research. While to our knowledge, pronominal verbs were never treated in a CBI fashion, the fact that several studies showed the success of this approach on the acquisition of grammatical concepts with a similar degree of complexity, such as German light-verb constructions (Suñer & Roche, 2019) or English modal verbs (Qin et al., 2023) is a testimony of the potential of CBI. Indeed, this instruction module would encompass several elements that we found essential to improve the way French pronominal verbs could be taught to Japanese university students in a thorough manner:

- 1) Pronominal verbs represent an abstract concept difficult to master for L2 learners, especially for those who lack corresponding reference in their L1 and even at intermediate levels of acquisition;
- 2) CLT instruction does not guarantee the acquisition of pronominal verbs in Japanese universities, as researchers report a general tendency, among learners, to avoid using them and leave them as dormant structures, even though the concept is prominently used in French;
- 3) Although it represents a challenging task, pronominal verbs can be explained through a SCOPA, likely because native speakers are able to recognise situations when a verb can be pronominalised.

To apply CBI as our experimental instruction project, we also needed to thoroughly explore the field of French pronominal verbs. In Section 2.2., we introduce the concept of pronominal verbs and the way it is described in the literature. The concept of diathesis is explained, as it plays a central role in our instruction module. A classification used as basis for our SCOPA is then constructed, notably with the annotation of a corpus of pronominal verbs.

2.2. Describing the target concept: pronominal-verbal constructions⁶

Providing a clear and exhaustive description of grammatical concepts is generally a serious challenge for linguists, although such descriptions have proven very useful in teaching situations, whether for native learners or for second and foreign language (L2) learners (Combettes, 2009). This challenge is of particular interest for the class of pronominal verbs in French, since this grammatical phenomenon is widely acknowledged as being especially complex to describe (Bień, 2007; Gaatone, 1975). Pedagogical grammars found in L2 textbooks – one of the most accessible resources containing classifications concerning French pronominals – have indeed been considered fragmentary and misleading regarding the way these verbs should be categorised (Connors & Ouellette, 1996; Delbarre, 2010). It is therefore not very surprising that pronominal verbs remain a major source of errors for L2 learners (Kraus, 1997; Nakamura, 2004), even though they are commonly used in all facets of our daily lives and conversations.

One widespread criticism of the grammatical descriptions of pronominal verbs is their lack of systematicity in the linguistic criteria used (Connors & Ouellette, 1996). While some categorisations of pronominal verbs emphasise morphosyntactic differences, such as the presence of a direct object, others are focused on semantic roles found in pronominal sentences (Bień, 2007). The difference in criteria implies that classifications of pronominal verbs can be difficult to compare. Moreover, several classifications tend to only categorise pronominal verbs at lemma level, disregarding their different meanings or uses, even though these verbs rarely have monolithic uses (Gaatone, 1975). These findings call for more thorough or more consistent

⁶ Section 2.2 is based on a publication (Zakhia Doueïhi & François, 2025)

classifications of French pronominal verbs, as this should facilitate their acquisition.

This section first aims to provide an overview of the linguistic classifications that have been designed to date, and further discuss their advantages and limits. We will then investigate a new type of classification of pronominal verbs based on the concept of diathesis. This concept, defined as the mapping of the constitutive syntactic functions onto the semantic roles in a verbal process, can be used methodologically as an anchor to understand and classify linguistic structures, including pronominal verbs, in a more efficient way. Systematic use of diathesis as a tool for analysis has the additional advantage of clarifying the relationship between the pronominal verb and its non-pronominal counterpart. Indeed, several authors (Bień, 2007; Corminboeuf & Heyna, 2015; Creissels, 2007; Dik, 1985) have postulated that the underlying coreferentiality within pronominal-verbal constructions reduces the valency⁷ of the verb, which means that the presence of *se* would necessarily have an impact on the diathetic mapping of the process when compared with its non-pronominal variant. As this hypothesis has not yet been empirically proven, this work will be the opportunity to verify it. Finally, approaching pronominal verbs as verbal structures based on a combination of semantic, syntactic, and valency analyses will also highlight their different uses in context, allowing us to systematically cover the polysemy of pronominal verbs.

Before moving on to the review of existing classifications of pronominal verbs in [Section 2.2.2.](#), we will first discuss some defining concepts about

⁷ Valency is used to describe the number of arguments a verb can potentially take; for instance, a divalent verb predicates a subject and a direct object. Valency can also be applied to semantic roles (Allerton, 1982).

pronominal verbs in the next section.

2.2.1. What is a pronominal verb?

According to Bouard (2011), the category of French "pronominal verbs" did not exist as such in grammar manuals before the early 18th century. This category used to be seen as a transformation of the active voice thanks to the clitic *se* and the auxiliary *être*, or as a particular case of neutral case. However, 18th century grammarians started to consider pronominal verbs as a distinct structure and attempted to define their properties, as their use was increasingly shifting away from their original reflexive meaning. This divergence might originate from the fact that in Old and Middle French, the so-called middle⁸ meaning (intransitive use of a verb describing an internal process) and the reflexive meaning (formed with *se* + verb) that were originally separate merged into a single form, as illustrated in the example in Table 8. This created greater ambiguity in modern French, as the same form might correspond to either the middle or the reflexive meaning.

	Middle meaning	Reflexive meaning
Old French	Il lieve [He wakes up]	Il se lieve [He stands up]
Modern French	Il se lève [He wakes up/He stands up]	

Table 8: Fusion of middle and reflexive meanings (Schøsler, 1999)

In modern French, a pronominal verb is made up, from the syntactic point of view, of a personal pronoun (*se*) in agreement with the grammatical subject, and a verb in the active voice (Connors & Ouellette, 1996; Joffre, 2018). In French, this formal structure is a prerequisite for the verb to be considered a pronominal verb.

⁸ Middle meaning derived from the Latin *media-tantum* forms, which are active intransitive verbs with a passive form.

From a semantic point of view, coreferentiality between the agent (A) and the patient (P) is sometimes seen as inherent to pronominal verbs (Gaatone, 1975). This implies that A and P should have no intrinsic semantic difference in order to acknowledge coreferentiality between the grammatical subject and the pronoun *se*, so that they could be exchanged without affecting the meaning of the sentence (Creissels, 2007). However, we can see that this is not systematic with verbs having more than one reading, such as *Je me bats*, which is commonly understood as [*I fight*] rather than its direct reflexive meaning [*I beat myself*].

Other than the presence of *se*, identifying common properties within pronominal verbs is not an easy task. Corminboeuf & Heyna (2015) suggest that the clitic *se* has the effect of reducing the valency of the verb, which means that it often prohibits the presence of either the agent (in the case of passive pronominal-verbal constructions) or the patient (in the case of direct reflexive sentences), as can be seen in the following examples:

	Subject	Verb	Object
Neutral active	Il [Agent]	endort	l'enfant. [Patient]
Agentive passive	L'enfant [Patient]	est endormi	par John. [Agent]
Passive pronominal	Un enfant [Patient]	s'endort (facilement).	
Direct reflexive	Il [Agent]	s'endort.	

Table 9: Valency reduction in pronominal sentences (Corminboeuf & Heyna, 2015)

For instance, we can see that in Table 9, the valency of *s'endormir* (to fall asleep) is reduced in comparison with *endormir* (to put to sleep). This valency reduction, operated through the presence of *se* could, according to the authors, be a better indicator of a pronominal construction. The role of *se* as a valency operator influencing the process in terms of agentivity is also hypothesised by Bień (2007) and Creissels (2007), although it has not been empirically proven in a large-scale study.

The observations made in this section allow us to define pronominal verbs as verbal structures taking the form NP *se* V, composed of a reflexive pronoun (*se*) agreeing with the grammatical subject (NP) and followed by a verb (V) (Burston, 1979). Throughout this chapter, we will attempt to prove, through a systematic methodology, that the presence of the reflexive pronoun does cause a reduction in valency of the process by one unit, for instance by dismissing the patient, the agent or the dative of a verbal process. The hypothesis that the clitic *se* modifies the verb's valency and, thus, the whole sentence from a diathetic perspective also implies that pronominal verbal constructions could be analysed according to their diathetic value. This analysis of the diatheses applicable through the use of pronominal verbs, also premised by Rivière (1997) and then by Dixon and Aikhenvald (2000) as promising grounds for analysis, will be covered in Section 2.2.4 of this study.

2.2.2. Review of the literature on classifications of pronominal verbs

As mentioned above, in modern French, pronominal verbs do not always express coreferentiality between the agent and the patient at the semantic level. This semantic confusion has so far prevented researchers from agreeing on a shared classification of pronominal verbs. This section reviews several classifications that have been introduced and aims to highlight their strengths and shortcomings.

One of the first classifications of pronominal verbs was proposed by Dangeau (1714-1717) and, more recently, reported and refined by Zribi-Hertz (1987). The pronominal verbs are here classified according to syntactic criteria into four categories: reflexive (*Pierre se lave*), reciprocal

(*Pierre et Marie s'aiment*), passive (*Le vin se boit frais*), and subjective (*Je m'évanouis*). In the reflexive derivation, the clitic represents a direct accusative object to the grammatical subject (*Pierre lave lui-même*); in the reciprocal derivation, the presence of the clitic implies that multiple subjects become, in turn, objects performing the process on each other (*Pierre aime Marie et Marie aime Pierre*); in the passive derivation, the clitic allows the subject to act as the object of a general statement in an active derivation (*On boit le vin frais*). This classification remains popular, as it is still used in some modern French textbooks and used by many researchers as a starting point for further analysis, thanks to its simplicity (Bień, 2007). Its main shortcoming lies in the heterogeneity of the last category (subjective), which encompasses all pronominal verbs that do not belong to the first three categories (Bień, 2007). Zribi-Hertz (1987) therefore attempted to further subdivide this category into the following classes: intransitive with unused non-pronominal counterpart (*s'évanouir*; *se repentir*), intransitive with non-reflexive counterpart (*se douter*; *se jouer*), intransitive with reflexive counterpart (*se mourir*), and intransitive with a similar meaning as the non-pronominal variant (*se casser*; *se gonfler*). While this subdivision provides a clearer view of the subjective category, the classification still has the disadvantage of considering many verbs as belonging to a single category, thus denying their polysemy, as well as not taking lexicalised pronominal-verbal structures into consideration (such as *se prendre la tête*).

By comparing syntactic structures from pronominals and their transitive variants, Boons, Guillet, and Leclère (1976) identified six types of relationships: reflexive (*Pierre se chatouille*), reciprocal (*Pierre et Marie s'aiment*), body parts constructions (*Pierre se lave les pieds*), possessive reflexive (*Pierre se dépense à mener à bien ce travail*), intransitive (*Le pont-*

levis s'est abaissé), and "phantom agent" constructions (*Le vin se boit frais*). In each case, the structure of the transitive variants of a verb is analysed before they are turned into their pronominal form, for instance with the transitive syntactic structures NP V N1 (*Je chatouille Marie*), NP V à N1 (*Je parle à Marie*) all turning into NP *se* V (*Je me chatouille/Je me parle*) to form the "reflexive" category. This classification has the advantage of using a methodology linking the pronominal structures with their transitive equivalent, making the meaning of the structures clearer (Bień, 2007). However, the focus on syntactic arguments at the expense of the semantic dimension is not without flaws, as it would, for instance, make it impossible to determine the structure of verbs that have no record of being used in a non-pronominal equivalent (François, 2011).

Geniušienė (1987) proposed a division between subjective and objective pronominal structures. Her classification is based on the recession of syntactic arguments: subjective structures are marked by object recession, while objective structures are marked by subject recession. A good example is the distinction between the autocausative sentence *Je me lève* [I wake up], in which the object is depleted (subjective structure), and the anticausative sentence *La porte s'ouvre* [The door opens], where the subject is depleted (as the action of opening cannot be performed without an external subject). Although this classification does not seem to consider structures where the recession is unclear (such as *Je me lave les dents*), the overall methodological soundness and the relevance of this distinction was praised by Melis (1990a), who proposed an adapted version of Geniušienė's model.

Melis's taxonomy (1990b) was based upon an experiment made with native speakers, who were confronted with made-up pronominal structures (such as

Il se pratèle, Ils se pratèlent and *ça se pratèle*) and asked to give an interpretation for each one of them. As a result, the researcher found out that along with the subjective structures and the objective structures already mentioned by Geniušienė (1987), there existed dative structures (logically created with the interpretations behind the made-up structure *Il se barufle les ouillais*). Melis argued that the distinction between structures was often made comparatively rather than intrinsically; as such, the complexity behind the interpretation of pronominal verbs could be tied to the understanding of valency and even aspect. This taxonomy notably puts the emphasis on contextual information and lexicon roots to provide better interpretations.

Following these works, Desclés and Guentchéva (1993) suggested a classification of pronominal verbs based on the interpretation of the roles of the agent and patient, and divided the structures into five categories (reflexive, reciprocal, middle, mediopassive, and passive). The breakdown follows the following pattern:

Structure	Example	Agent	Patient
<i>Reflexive</i>	<i>Pierre se rase</i>	Pierre	Pierre
<i>Reciprocal</i>	<i>Les enfants se battent</i>	Les enfants	Les enfants
<i>Middle</i>	<i>Le vautour s'abattit sur le lapin</i>	Le vautour	Le lapin
<i>Mediopassive</i>	<i>Le mât s'abattit avec un grand fracas</i>	Ø	Le mât
<i>Passive</i>	<i>Les livres se vendent bien</i>	"Les gens"	Les livres

Table 10: Desclés and Guentchéva's classification (1993)

The main strength of this classification is that it sheds light on the identification of specific semantic roles as a major distinctive marker of pronominal structures (Bień, 2007).

Noticing the huge differences between the nature of criteria chosen to classify pronominals in the previous typologies, Bień (2007) mixed together

the most prominent distinctive features among them (coreferentiality, syntactic recession, syntactic transitivity, agentivity, and verb lexicality) and suggested a semantic classification of pronominal verbs based on two criteria: agentivity and control over the process, as can be seen in Table 11.

<i>Structure</i>	<i>Example</i>	<i>Agentivity</i>	<i>Control</i>
<i>Active</i>	<i>Jean se lave</i>	+	+
<i>Medio-active</i>	<i>Jean s'énervé</i>	+	0
<i>Mediopassive</i>	<i>Le mât s'écroule</i>	0	0
<i>Passive</i>	<i>Le problème se discute</i>	0	-

Table 11: Bień's (2007) classification

Bień (2007) describes the agentivity feature as the grammatical subject acting on the process as an agent, while the control feature would be the "*capacity to carry out an action and capacity to interrupt it*". This classification enables the well-known concept of agentivity to be divided between two sub-concepts (agentivity and control) that are semantically related, but not equivalent. It also has the merit of allowing lexicalised *se*-verbs (such as *s'évanouir*) to be analysed as medio-active structures (Bień, 2007). However, other types of structures are indiscriminately mixed. For instance, the structure *Ils se lavent* is only categorised as active, whereas it can have at least two different semantic interpretations: [*they wash themselves*]/[*they wash each other*].

Following the tenets of construction grammar (Goldberg, 1995), Petersilka (2012) classified pronominal structures based on both syntactic and semantic criteria, as well as on whether the root of the pronominal verb is a transitive verb. The analysis resulted in a 10 classes distribution: attribute-value (*Tu t'es trahi*), reflexiva tantum (*Elle s'est évanouie*), anticausative (*Le soleil s'est couché*), mediopassive (*Cette tour se voit de loin*), reciprocal (*Pierre et Jean se trompent*), reciprocal reflexiva tantum (*Les gamins se chamaillent*

sur la place), reciprocal anticausative (*Nos routes se séparent ici*), completive (*On s'attendait au pire*), attributive (*Ils se pensent souvent drôles*), and attributive with thing-object (*Cette école se veut complémentaire d'autres initiatives*). Methodologically speaking, the way this typology takes syntax, semantics, and even etymology into account is commendable; however, the methodology behind the choice of the sample structures is not easy to figure out, leading to some degree of confusion. Consequently, it is possible that some existing structures do not fall into any of these categories. Additionally, some categories are mainly differentiated by semantics, such as completives (e.g. *On s'attendait au pire*), do not attempt to describe how semantics differ with pronominalisation (e.g. there is no information on how *On attendait* differs from *On s'attendait*), merely claiming that they take a different meaning. Hence, the classification could benefit from a more thorough examination of the classes it constructed.

By using systematic corpus analysis based on formal criteria, Barque, Candito and Huyghe (2018) built a typology of pronominal verbs. Through the annotation of pronominal sentences in two corpora, **Sequoia** (318 tokens) and **French Treebank** (or FTB, 900 tokens), the researchers were able to identify 15 categories of pronominal structures: intrinsic, autonomous (subdivided into five categories: syntax, selection, aspect, lexicalised, and construction), autocausative, reflexive, reciprocal (including two types: V and *se*-V), mediopassive, anticausative, antipassive, parallel, and autobenefactive. The annotation process showed good results in terms of annotator agreement (Kappa values range from 0.53 – with 13 categories involved – to 0.74 with a broader granularity), although it emphasised how difficult it is to distinguish reflexive interpretations based solely on formal criteria. All in all, the typology appears to be complexified by the large

number of categories. However, the use of corpus annotation to certify the typology is particularly innovative.

As can be seen, there have been attempts to classify pronominal verbs according to different principles (syntactic, semantic, or even etymological). However, it sometimes remains difficult to determine whether these authors classify similar verbs in the same way (e.g. whether Zribi-Hertz's passive category is equivalent to Boons, Guillet and Leclère's "phantom agent" constructions). To have a better idea of how these classifications compare with each other, we sampled 20 verbs from all categories proposed by Zribi-Hertz (1987) and attempted to classify the most frequent uses of these verbs under each classification (see Annex 1).

The data collected show that while some structures can easily be agreed upon, other verbs such as *s'évanouir* are difficult to classify (it has seven different designations in the seven classifications). Additionally, it appears that earlier typologies rarely take the polysemy of pronominal verbs into account, and the overall tendency to analyse semantic and syntactic parameters separately can lead to many exceptions. Moreover, the sample coverage is limited and rarely supported by corpus analyses (only found with Barque et al., 2018), leading to potential methodological weaknesses. To overcome these limitations, we wish to suggest a new classification of French pronominal verbs based on the concept of diathesis. By means of a two-phase methodology combining systematic analyses and corpus annotation, we aim to build a typology with enough clarity for potential educational purposes, while remaining precise and thorough enough for linguistic depictions. Before presenting this new typology and the results of the annotation, the next sections will briefly summarise the notion of

diathesis and how it applies to pronominal verbs, as this concept is essential in asserting our taxonomy.

2.2.3. Application of diathesis to pronominal verbs

2.2.3.1. About the concept of diathesis

The concept of diathesis in morphosyntax can be defined as "*a pattern of mapping of semantic arguments onto syntactic functions (grammatical relations) [...]*" (Kulikov, 2011). According to Kulikov (2011), the most common mapping is "X = S; V; Y = DO", where the first semantic role (X) is mapped onto the syntactic function of *subject* (S), and the second semantic role (Y) is mapped as *direct object* (DO). This diathesis, the most common one, is often regarded as the active or neutral diathesis (Rousseau, 2014; Shibatani, 2006). Changing the mapping of the arguments alters the diathesis and can impact the valency of the verb. For instance, promoting the direct object of a neutral diathesis to subject and demoting the original subject to an oblique object (or removing it) will result in a change of diathesis, as seen in the following example:

1. I see an apple.
2. An apple is seen (by me).

Diatheses are often presented in tabular form for the sake of clarity and convenience. The following table (Table 12) corresponds to the above example of change of diathesis for English:

<i>Example</i>	<i>I</i>	<i>see</i>	<i>an apple</i>
<i>Semantic role</i>	A	V	P
<i>Syntactic function</i>	S	V	DO
↓			
<i>Example</i>	<i>An apple</i>	<i>is seen</i>	<i>(by me)</i>
<i>Semantic role</i>	P	V	A
<i>Syntactic function</i>	S	V	Ø (OBL)

Table 12: Change from neutral to passive (agentless) diathesis

Since changing a syntactic function or a semantic role is a sufficient condition to trigger a switch of diathesis, there are many diatheses in every language. A quick look at the following list of the main syntactic functions and semantic roles (hierarchically ordered according to Grimshaw, 1990) reveals a high number of potential arguments that could be combined into a diathesis.

Syntactic function	Explanation	Example
Subject (S)	Object predicated by the verb in the clause	<i>Jean mange la pomme.</i>
Direct Object (DO)	Object linked directly to a transitive verb	<i>Jean mange <u>la pomme</u>.</i>
Indirect Object (IO)	Object linked by a preposition to a transitive verb	<i>Jean offre un cadeau <u>à Marie</u>.</i>
Oblique Object (OBL)	Object linked by a preposition to an intransitive verb	<i>Jean joue <u>au poker</u>.</i>
Dative (DAT)	Object towards which the verb is directed in the clause	<i>Marc écrit <u>pour sa mère</u>.</i>
Genitive (GEN)	Object attributive to another object	<i>Les oreilles <u>du chat</u> sont petites.</i>
Adjunct (ADJ)	Non-mandatory object	<i>Jean mange la pomme <u>sur un banc</u>.</i>
Dummy (∇)	Dummy function (cf. impersonal <i>Il, ça</i>)	<i><u>Il</u> pleut.</i>
Semantic role	Explanation	Example
Agent (A)	Entity performing the process	<i>Jean mange la pomme.</i>
Patient (P)/ Undergoer (U)	Entity undergoing the process	<i>Jean mange <u>la pomme</u>.</i>
Beneficiary (B)	Entity benefiting from the process (positively or negatively)	<i>Jean offre un cadeau <u>à Marie</u>.</i>
Experiencer (E)	Entity experiencing the process (state of being, emotion, perception)	<i>Jean a peur.</i>
Instigator (I)/ Causer (CAUS)	Animate entity causing the process to happen	<i>Marc demande à Jean de venir.</i>
Cause (CAUS)/ Stimulus (STI)	Inanimate entity causing the process to happen	<i><u>La pluie</u> a fait annuler la représentation.</i>
Circumstant (C) [Locative, Modal, Temporal...]	Entity circumstant to the process	<i>Jean mange la pomme <u>sur un banc</u>.</i>
Dummy ($\emptyset$)	Does not refer to any entity	<i><u>Il</u> pleut.</i>

Table 13a & 13b: Main syntactic functions and semantic roles in French

Despite the large number of possible combinations, some diatheses are more prominent in a language. Table 14 shows some major diatheses used in French in Kulikov's (2011) typology.

Diathesis	Example	Syntactic mapping	Semantic mapping
<i>Neutral</i>	Je mange une pomme.	S V DO	A V P
<i>(Agentless) passive</i>	La pomme est mangée (par quelqu'un).	S V	P V (A)
<i>Impersonal</i>	Il pleut.	∇ V	$\emptyset$ V
<i>Canonical reflexive</i>	John se rase.	S V	A = P V
<i>Canonical reciprocal</i>	Pierre et Marie s'embrassent.	S V	A & P V
<i>Anticausative</i>	La branche casse.	S V	P V
<i>Indirect causative</i>	Je lui ai fait manger la pomme.	S IO V DO	I A V P

Table 14: Major diatheses in French

The central role of the neutral diathesis (*Je mange une pomme*) is justified by the fact that many diatheses can be considered extensions of it, such as the indirect causative (*Je lui ai fait manger une pomme* – [I made him eat an apple]).

2.2.3.2. Diathesis with pronominal verbs

French pronominal verbs are frequently cited as an example of diathetic change from the neutral use of the verb (Deronne, 2006; Bień, 2007; Weidenbusch, 2016). As we saw in Section 2.2.1, pronominal verbs have an impact on the valency of the verb, as they affect the semantic and syntactic arguments of the sentence, and they could be represented by one or more diatheses. In this regard, Kulikov's (2011) typology of diatheses based on the valency of the arguments provides useful insights into which ones could

apply to pronominal verbs. Derivative diatheses (such as various passives) preserve the semantic arguments of agent and patient while demoting them syntactically. On the other hand, operational diatheses (such as reflexive and reciprocal) keep the original meaning of the arguments taken from a neutral derivation, but add some operation to them (such as coreferentiality). Finally, diatheses that add or remove semantic arguments include causative, autocausative, applicative, benefactive, and anticausative clauses. While the names given to these diatheses may vary, the patterns they describe remain the same in studies on morphosyntax (Dixon & Aikhenvald, 2000; Le Querler, 2012; Poole, 2014; Weidenbusch, 2016).

The interest of considering valency changes when comparing neutral diathesis with other structures, as shown in Kulikov's typology (2011), is that pronominal verbs, for instance, are known to have an impact on valency (Corminboeuf & Heyna, 2015). Thus, to determine which diatheses are covered by pronominal verbs, observing how the valency of the verbal process changes could be key. Indeed, although some researchers have hinted at the potential usefulness of the concepts of diathesis and valency regarding pronominal verbs (Bień, 2007; Creissels, 2007; Corminboeuf & Heyna, 2015), no systematic effort has yet been made to apply these concepts to a classification of pronominal-verbal constructions based on semantic, syntactic, and empirical criteria.

Hereafter, we will aim to determine whether a diathetic classification of pronominal structures – rather than forms – integrating semantic, syntactic, and empirical analyses all together would be possible, and how valency changes occur within pronominal-verbal constructions. To achieve this aim, Kulikov's (2011) typology of diatheses will be used as the main basis of our

analysis. The next section will describe the methodology used to build our classification of pronominal verbs based on diathesis.

2.2.4. Two-step analysis

To carry out our classification of pronominal verbs, we applied a two-phase procedure. The objective of the first phase was to determine which diathetic patterns would correspond to a given number of prototypical pronominal structures, while the second phase would attempt to verify that these diatheses could effectively cover all forms found in a large corpus.

2.2.4.1. First phase:

Identification of diatheses applicable to pronominal verbs

The first phase followed a three-step process: 1) selecting a representative set of verbs that will be subjected to analysis, 2) identifying the major uses for each chosen verb, and 3) determining adequate classification criteria.

Our first step was to determine a set of pronominal verbs representative of the different categories found in the literature (see Section 2.2.2). We took Zribi-Hertz's (1987) classification as a starting point for our analysis, as it is a modern approach inspired by the authoritative figure of Dangeau's typology in the field of pronominal verbs (Bień, 2007). Indeed, Dangeau's work has been cited as the foundation of the “**quatripartition traditionnelle des verbes pronominaux**”, which still has prominent impact in grammar textbooks. Compared to other typologies we've reviewed in this paper, Zribi-Hertz chooses, in her own words, to keep the classical typology untouched apart from delving deeper into understanding of the subjective category. From there, we sampled up to three verbs per category:

Reflexive: *se laver, se regarder, se raser*

Reciprocal: *s'aimer, se voir, se parler*

Passive: *se boire, se vendre, s'utiliser*

Subjective, divided into:

- intransitive with unused non-pronominal counterpart: *s'évanouir, se repentir*
- intransitive with non-reflexive counterpart: *se douter, se jouer*⁹
- intransitive with reflexive counterpart: *se mourir*¹⁰
- intransitive with similar meaning as the non-pronominal variant: *se casser, se gonfler*.

We also supplemented these 16 verbs with additional pronominal structures in French, which, albeit rare, could not be ignored due to their unique structure, namely the impersonal pronominal verbs. This category of pronominal verbs was either absent from the classifications (Boons et al., 1976; Geniušienė, 1987; Bień, 2007) or mentioned by several authors to be referred to as part of the passive (Desclés & Guentchéva, 1993; Albrecht, 1997; Weidenbusch, 2016) or the subjective (Zribi-Hertz, 1982; 1987). We thus included four additional verbs chosen on the basis of an analysis of impersonal pronominal verbs carried out by Rivière (2018) (*s'agir, se passer, se construire, se pouvoir*). This enabled us to reach a total of 20 sample verbs.

As a second step, we searched for typical uses of each of these 20 verbs in their dictionary entry in the Trésor de la Langue française informatisé (TLFi). The TLFi is one of the most exhaustive online dictionaries for French, with over 100,000 entries, and is useful for its historical references. Based on this source, we were able to identify 33 different structures for our

⁹ *Douter* and *jouer* are both intransitive.

¹⁰ *Mourir* has a reflexive meaning.

20 verbs, in authentic and historical contexts, which is important given that diathetic mapping should be analysed in context.

Our third step was to determine the classification criteria on the basis of which the various diathetic mappings found in our 33 structures would be described. The four criteria we used for the analysis are based on Kulikov's (2011) typology: the syntactic functions of the verb arguments, their semantic roles, their cases, and valency changes compared with the neutral diathesis. Each of these criteria is described in more detail below.

First criterion: syntactic function

The syntactic functions of the arguments of the pronominal verb are one of the two essential characteristics in diathetic mapping. The nature of a syntactic function is determined by the function an argument takes in the verbal process. Explanations of these functions can be found in Table 13a & 13b.

Categorical (nominal): subject (S), direct object (DO), indirect object (IO), oblique object (OBL), dative (DAT), genitive (GEN), adjunct (ADJ), verb (V), not applicable (Ø).

Second criterion: semantic role

The semantic role of the arguments is the other essential marking in diathetic grammar. The nature of a semantic role is determined by the meaning an argument takes in the verbal process. Details concerning semantic roles can be found in Table 13a & 13b.

Categorical (nominal): agent (A), patient (P), experiencer (E), beneficiary (B), instigator (I), cause (CAUS), circumstant (C), verb (V), not applicable (Ø).

Third criterion: case-marking

Case-marking is a grammatical feature indicative of the syntactic function of the arguments. This criterion is often coded in conjunction with the syntactic function for the sake of clarity, as it may give a more flexible understanding of syntactic functions¹¹.

Categorical (nominal): nominative (NOM), accusative (ACC), dative (DAT), genitive (GEN), ablative (ABL), verb (V), not applicable (Ø).

Fourth criterion: valency change in comparison with the neutral diathesis

The valency change compared with the neutral diathesis is essential to determine internal properties of pronominal verbs. The transformation into the neutral diathesis will be provided by turning the arguments into an "X = S; V; Y = DO" mapping, while keeping the meaning of the arguments (semantic roles) as close to the pronominal variant's arguments as possible. In cases where the pronominal sentence would not have a grammatically correct neutral diathesis counterpart, the clitic *se* will be replaced by the object argument (direct or indirect) *soi-même* [oneself], as it is regarded as both the historical meaning of the reflexive pronoun and the closest substitute for it. The method to determine whether the valency has decreased or increased will be based on Grimshaw's (1990) argument structure theory. Categorical (nominal): addition, reduction, promotion, demotion (of an argument), fusion (of arguments)

These criteria will be coded in tabular form, commonly used for diathetic mapping representations. Table 15 is an example of the format of a diathetic table:

¹¹ In the sentence "*Il s'achète un cadeau*", the interpretation of *se* can either be IO, possessive, or even ADJ; however, the case designation of Dative encompasses all these functions in one category.

	<i>Je</i>	<i>me lave</i>
<i>Syntactic function</i>	S	V
<i>Semantic role</i>	A	V
<i>Case-marking</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: * <i>Je lave moi-même</i> ► Reduction of a syntactic argument (V cannot take DO) ► Reduction of a semantic argument (V cannot take P)	

Table 15: Example of diathetic mapping

2.2.4.2. Classification of pronominal verbs based on diatheses

A general look at the diatheses covered by the sample verbs

As a result of the analysis of the diatheses corresponding to the 33 different structures analysed, our inductive approach identified eight different mappings of diathesis. These are broken down in the following table, along with the three types of properties we analysed, i.e. syntactic, semantic, and empirical.

Diathesis	Syntactic	Semantic	Empirical	Examples
Canonical reflexive	S se V	A se V	A = P Valency decrease ($\text{D}\Theta/\text{P}$)	<i>Je me lave</i>
Canonical reciprocal	S se V	A se V	A α = P α Valency decrease ($\text{D}\Theta/\text{P}$)	<i>Pierre et Marie s'aiment</i>
Reflexive benefactive	S se V DO	A se V P	A = B Valency decrease ($\text{D}\Delta\text{F}$)	<i>Le chat se lave l'oreille</i>
Autocausative	S se V	E se V	E = CAUS Valency decrease ($\text{D}\Theta$)	<i>Je me suis calmé</i>

Potential (agentless) passive	S se V	P se V (A)	A = ? Valency decrease ($\emptyset\emptyset$)	<i>Le français se parle dans ce village</i>
Conversive passive	S se V OBL/ADJ	P se V A	A = OBL/ADJ Valency decrease ($\emptyset\emptyset$)	<i>L'avion se remplit avec les passagers</i>
Impersonal passive (w/ dummy subject)	∇ se V DO/OBL	$\emptyset$ se V P	S/A = $\emptyset$ Valency decrease (S/A)	<i>Il se passe des choses étranges</i>
Anticausative	S se V	P se V	A/CAUS = $\emptyset$ Valency decrease ($\emptyset\emptyset$)	<i>Le verre se cassa</i>

Table 16: Typology of diatheses represented by pronominal verbs

From a syntactic point of view, most of these constructions follow the structure S *se* V, with only a few exceptions (reflexive benefactive, conversive, and impersonal passive). While the structure S *se* V only represents the minimal syntactic mapping of a given structure, this relatively common pattern for pronominal verbs could explain the difficulties in interpreting each type of construction.

Semantically, Table 16 distinguishes between three major categories of diatheses represented in pronominal-verbal constructions:

- Diatheses where the subject of the sentence is preserved, but an external argument is demoted/removed (i.e. canonical reflexive, reciprocal, reflexive benefactive, anticausative).
- Diatheses where a patient is promoted as the main predicate and the agent is demoted/removed (i.e. potential passive, conversive passive, impersonal

passive, anticausative).

-Diatheses where the process is not predicated by an agent/patient dynamic (i.e. anticausative).

Empirical observations will supplement syntactic and semantic mappings by linking them together, as well as describe operations on the valency of the arguments. A quick analysis of the constructions enables us to pinpoint the following specific features:

-The function of the reflexive morpheme in the canonical reflexive is to encode coreferentiality between the agent and the patient (they refer to the exact same entity).

-The function of the reflexive morpheme in the canonical reciprocal is to encode coreferentiality between at least two agents and patients that mutually perform the process.

-The function of the reflexive morpheme in the reflexive benefactive is to encode the addition of a dative nuance to the process that the agent carries out on the patient.¹²

-The function of the reflexive morpheme in the anticausative is to encode a self-caused sensorial process affecting the experiencer.

-The function of the reflexive morpheme in the potential (agentless) passive is to encode a modal process (possibility/obligation) with suppressed agent.

-The function of the reflexive morpheme in the conversive passive is to encode the demotion of the agent to a circumstantial position.

-The function of the reflexive morpheme in the impersonal passive is to encode the marginalisation of the subject/agent as a dummy subject.

-The function of the reflexive morpheme in the anticausative is to encode the

¹² While most of the time, this diathesis implies partial coreferentiality between the agent and the patient (such as an agent acting upon a part of itself in body-part constructions like *Je me lave les mains*), this is not a prerequisite condition (see below).

deletion of any cause and/or agent identifiable in the process.

2.2.4.3. Second phase:

Validation through corpus analyses of Sequoia and FTB

As a result of the first phase of our research, we identified eight diathetic patterns, which cover the 33 prototypical reflexive structures found in previous work. To verify the degree of applicability of these patterns and the potential need for additional diathetic categories, we carried out a manual corpus analysis of pronominal structures found in the *Sequoia* (332 occurrences) and the *FTB* (900 occurrences) corpora. *Sequoia* is composed of over 3,200 sentences sourced from *Europarl* (European Parliament texts), *L'Est Républicain* (a French right-wing newspaper), *Wikipedia*, as well as *European Medicines Agency*, while *FTB* includes 1 million words taken from *Le Monde* (1989-1995), one of the most prominent daily newspapers in France. Occurrences were identified by isolating the reflexive clitic *se* and its derivatives, preceded or followed by a verbal form. Three linguists (A1, A2, A3), among which two of them (A1 & A2) had no prior expertise with the concept of diathesis, were asked to determine the diathesis which each structure belongs to and to detail their reasoning, using a three-step procedure described in an annotation guide: 1) identifying the structure's meaning, 2) constructing a diathetic mapping, and 3) answering binary questions to interpretate the mapping. These steps are further detailed in the annotation guide (Annex 7). The annotation procedure was carried out in an iterative way: about 100 structures were first annotated by the three linguists (referred to as annotators from here on), leading to a discussion to resolve disagreements and produce a gold standard. The annotation guide was then refined, and a new 100-structure annotation phase was started. At the end of this second phase, we computed a Cohen's Kappa to assess **inter-rater**

agreement on the 200 structures. The values of κ are 0.37 for A1-A2, 0.44 for A1-A3, and 0.64 for A2-A3. It appeared that A1 had more trouble adapting to the typology, while A2 and A3 reached a fairly high degree of agreement. The 1,032 remaining structures were then separately annotated by the annotators after receiving final feedback. At the end of the annotation process, we defined a gold standard based on a consensus between the annotators.

Distributions of the categories

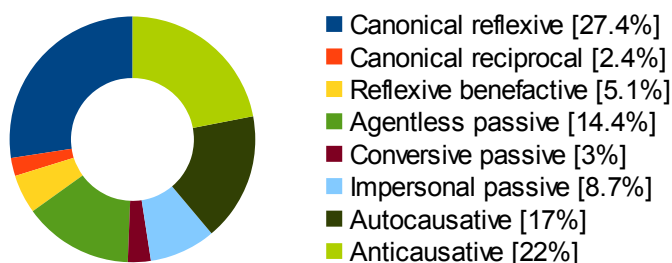


Figure 6a: Diathesis distribution in Sequoia

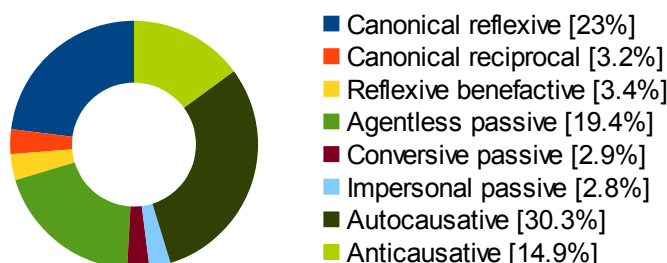


Figure 6b: Diathesis distribution in FTB

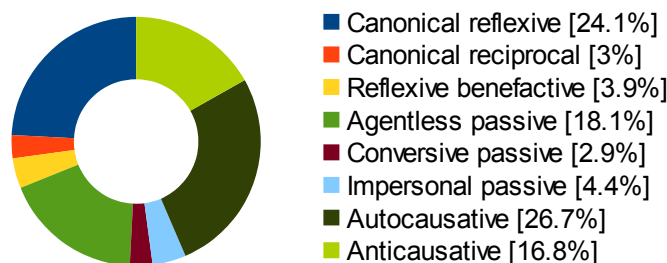


Figure 6c: Diathesis distribution in combined corpora

The first and most important result is that the annotation process and the discussions that followed did not highlight any new structures that did not support our typology. Our typology was then validated by this corpus analysis. The distribution of our eight categories in the corpus is set out in Figure 6a (*Sequoia*), Figure 6b (*FTB*) and Figure 6c (Combined corpora).

By adding up numbers from the two corpora, we can observe that four categories were especially well-represented: autocausive (329 structures = 26.7%), canonical reflexive (298 structures = 24.1%), potential agentless passive (223 structures = 18.1%), and anticausative (207 structures = 16.8%). However, the category distribution in the two corpora is significantly different ($X^2(7) = 51$; $p < 0.001$). Further analysis of these differences through adjusted standardised residuals shows that impersonal passives ($z = 4.53$) and anticausatives ($z = 2.95$) are overrepresented in *Sequoia*, whereas agentless passives ($z = 2.01$) and autocausatives ($z = 4.74$) are overrepresented in *FTB*. The genre of the texts forming these corpora seems to have considerable impact on the distribution of diathesis types. Autocausatives, for instance, are more common in general newspapers like *Le Monde/FTB* (273 occurrences), whereas medical literature generates a greater number of anticausatives, such as the following examples (*Sequoia*):

- (1) *L'acide zolédronique ne se lie pas fortement aux protéines.*
- (2) *Les os deviennent plus fragiles et se cassent plus facilement.*

As a general tendency, however, we can see that reciprocals, reflexive benefactives, and converse passives have low representation, potential agentless passives and impersonal passives have low to middle representation, and canonical reflexives, autocausatives, and anticausatives have high representation.

However, these figures are very much related to the definition of each class in our typology. By way of comparison, Table 17 compares our annotation with that of Barque et al. (2018), who divided the majority of the 900 FTB tokens into three categories: 372 various autonomous structures (41.3%), 221 anticausative structures (24.6%), and 104 reflexive structures (11.6%):

Category	Reflexive	Reciprocal (any kind)	Intrinsic	Autocausative	Anticausative	Medio-passive	Medio-passive anticausative	Autonomous (selection)	Autonomous (syntax)	Autonomous (lexicalised)	Autonomous (construction)	Autonomous (aspect)
Diathesis												
Canonical reflexive	57	4	5	38	8	5	0	52	27	7	3	5
Canonical reciprocal	2	24	0	0	0	0	0	2	0	1	0	0
Reflexive benefactive	12	2	1	1	2	0	1	2	2	4	0	0
Agentless passive	11	4	0	6	58	27	6	29	22	1	7	4
Conversive passive	2	0	0	1	2	10	0	11	0	0	0	0
Impersonal passive	0	0	0	0	0	0	0	3	22	0	0	0
Autocausative	18	0	25	21	60	4	2	53	85	4	0	1
Anticausative	2	2	4	3	91	3	4	13	3	0	0	9

Table 17: Distribution of pronominal structures in FTB, comparison with Barque et al. (2018)

As can be seen in Table 17, only 70 structures were identified as autocausatives by Barque et al. (2018), compared with 273 in our typology. However, a good number of autonomous came under our autocausative category, and we categorised some autonomous in Barque et al. (2018) as impersonal passives or potential agentless passives. While the definition of autocausative differs between our typologies, the huge gap between categories speaks for the difficulty of identifying some structures. Meanwhile, the number of reciprocals was close enough (36 to 29) to conclude that this category was rather easy to identify. Barque et al. (2018) also note that there were no occurrences of autobenefactives, parallels, and/or antipassives. Indeed, the structures we annotated as reflexive

benefactive did not correspond to these authors' autobenefactives, which were mostly considered as reflexives.

2.2.5. Discussion about polyfunctional categories

Now that we have an idea of the proportion of each diathesis within two French corpora, we need to look at the difficulties that occurred during the annotation process to assert the reliability and simplicity of our typology. As such, we computed the **recall** or sensitivity between each annotator and the gold standard (GS) - a value of 1 for a given diathesis means that 100% of GS annotations were correctly identified by the annotator -, and the **true positive rate** (true pos) or positive prediction value - a value of 1 for a given diathesis means that all occurrences assigned to this diathesis were correct. The recall rate and true positive rate of each diathesis between the annotators (A1, A2, and A3) and the gold standard (GS) in Sequoia (1-200) are as follows:

Diathesis	Recall rate (A1 + GS)	True pos rate (A1 + GS)	Recall rate (A2 + GS)	True pos rate (A2 + GS)	Recall rate (A3 + GS)	True pos rate (A3 + GS)
Canonical reflexive	0.8333	0.5195	0.8125	0.7500	0.9167	0.8302
Canonical reciprocal	0.8333	0.5556	0.8333	1	0.8333	0.5556
Reflexive benefactive	0.1429	0.2222	1	0.8750	1	0.9333
Potential (Agentless) passive	0.4348	0.3704	0.9130	0.7500	0.7826	0.8571
Conversive passive	0.2857	0.2000	0.4286	0.5000	0.8571	1
Impersonal passive	1	1	1	0.8947	1	1

Autocausative	0.4324	0.4000	0.5135	0.7600	0.9459	0.8140
Anticausative	0.08333	0.36364	0.8542	0.8367	0.7083	0.9444

Table 18: Confusion matrix recall and true positive rates compared with the gold standard, Sequoia 1-200

Noticeably, some categories were more prone to mistakes than others. A1 and A2 both had trouble with the conversive passive, given their low to average scores in both recall and true positive rates. The autocausative was also difficult to identify for A1 & A2, although A2 rarely assigned this diathesis incorrectly. Individually, A1 had difficulties in both identifying and distinguishing the reflexive benefactive and the anticausative diatheses. On the bright side, the canonical reflexive, reciprocal, and impersonal passive diatheses obtained very high rates for each annotator. The agentless passive was also fairly well defined.

Consequently, four diatheses, namely the reflexive benefactive, the conversive passive, the autocausative, and the anticausative diatheses were seemingly not clear enough for annotators, especially A1. We will hereafter describe the difficulties relating to their interpretation and the rationale for our classification.

The case of reflexive benefactive

This diathesis is defined by two major characteristics: the necessary presence of an object (most of the time, a DO/P) and a valency reduction operated through the suppression of a dative. Evidence of this dative can be found in the neutral diathesis through the addition of an IO, a possessive, or a

genitive, which all take the semantic role of beneficiary:

(3) *Il découvre un lion qui se frotte le ventre.*

→ *Un lion qui frotte le ventre à/pour lui-même* (IO)

→ *Un lion qui frotte son ventre* (possessive)

→ *Un lion qui frotte le ventre de lui-même* (genitive)

In each sentence, the semantic interpretation is that of a process aimed at the agent, which implies it also becomes a beneficiary.

This interpretation of the agent/beneficiary duality is supported by Creissels (2007) and can also be applied to a sentence such as *Je me mange tout le gâteau*. Here, the agent performs the process (manger) on the patient (le gâteau) to produce an effect on himself (emphasised by the clitic). Typically, eating involves the participation of the agent in the process to a higher degree than other actions, since it is most often motivated by a goal (satiety, for instance). With this in mind, it is not difficult to understand why the agent would also be a beneficiary in a sentence such as *Je me mange tout le gâteau*. However, this effect can be involuntary and detrimental, as can be seen in the sentence *Je me prends un coup*. *Se prendre* should be understood here as meaning [*to take for oneself*]. A lot of words can be combined with this meaning of *se prendre* and will always take either a positive, voluntary nuance (known as real benefactives), or a negative, involuntary nuance (known as malefactive¹³):

(4) *Je me suis pris un hamburger.* (benefactive)

13 Zúñiga & Kittilä (2010) offer a detailed perspective about the nuances of benefactive and malefactive in their reports, especially in sentences with *give* and *take*

(5) *Je me suis pris une boisson.* (benefactive)

(6) *Je me suis pris une amende.* (malefactive)

(7) *Je me suis pris une gifle.* (malefactive)

Although the identity of the agent in such sentences could be debatable due to the semantic construction, from a syntactic point of view, the action of *se prendre* is still carried out by the subject *Je*, which is why we would categorise these sentences as reflexive benefactive. Indeed, they do share the same diathetic mapping (A *se* V P - S *se* V DO), and we were also able to recover these benefactive/malefactive nuances in other types of sentences categorised under this diathesis:

(8) *Je me lave les mains.* (benefactive)

(9) *Les patients qui se sont fracturé la hanche.* (malefactive)

(10) *Je me tape les bouchons.* (malefactive)

(11) *Je me réserve le droit de prendre le volant.* (benefactive)

All these sentences belong to the category of dative turns as defined by Evrard (2003), since they revolve around the same principle of an agent being treated as a beneficiary/maleficiary. Although there were few occurrences of reflexive benefactive in the corpus analysis (respectively 17 & 31 tokens, respectively, totaling 3.8%) in the corpus analysis, it should be easy to identify thanks to its peculiar structure. According to the feedback from our annotators, the confusion with other types came from misidentifying the dative meaning in specific sentences.

The case of conversive passive

Conversive passives, much like the agentless passive, are diatheses placing

the patient in subject position. However, unlike the agentless passive, this diathesis allows the neutral diathesis agent to be expressed, although in a marginal position (adjunct or oblique complement):

(12) *L'avion s'est rempli avec les passagers.*

(13) *La finale se jouera entre Utrecht et Soleuvre.*

(14) *La séance s'est terminée par des applaudissements.*

There were very few occurrences of this diathesis in the corpus (respectively 10 & 26 tokens, respectively, totaling 2.9%), but it was often confused with aspectual anticausatives, such as (14). The difference between conversive passives and aspectual anticausatives is mainly found in the semantic nature of the adjunct/oblique complement. For instance, unlike *les passagers* towards *remplir* in (12), *des applaudissements* cannot be the active agent of *terminer*.

The case of autocausative

Unlike reflexives which license dynamic processes controlled by an agent, autocausatives are defined as events in which a conscious being is subject to self-induced sensorial experience (Parry, 1998). As such, prototypical examples of the autocausative diathesis can be found within emotion verbs (e.g. *s'énerv*, *s'attrister*, *se fâcher*, *se réjouir*, *s'ennuyer*, *s'effrayer*, *s'étonner*, *s'honorer*, etc.). Pronominal structures are particularly suitable for expressing emotions, as the semantic role of experiencer is prominent in such constructions (Croft, 1991). In a sentence such as *Je m'énerv*, *je* is hardly an agent; otherwise, it would imply that *je* is in total control of the process. The fact that autocausatives license experiencers instead of agents/patients is also reflected in **tests of patient emphasis** such as *Je m'énerv moi-même**,

which is less acceptable than the emphasised canonical reflexive *Je me lave moi-même* (de Schepper, 2010).

Similarly, Kailuwet (2007) writes that emotion verbs in French turn the subject into an experiencer, especially because the degree of controlled intentionality is lower than in action-focused verbal processes. The same analysis was made by Bień (2007), who treats these verbs as medio-active pronominal structures (agentivity +, control 0). Meanwhile, the diminished control over the process (due to the nature of inchoative verbs) does not prevent the experience from happening willingly either. As such, in (15) and (16), the subject can be considered an experiencer because the verb describes a voluntarily self-induced change of state:

(15) *Le titulaire de l'AMM s'engage à fournir des études.*

(16) *Le Parlement s'est efforcé d'en tenir compte [...]*

Other inchoative verbs potentially characterised as autocausatives would include physical reactions (*se réveiller, se fatiguer, s'épuiser, se mourir*) or perceptive and mental attitudes (*se repentir, s'apercevoir, se douter*), due to the fact that these verbs feature change of state/perceptive events within a conscious experiencer, rather than controlled actions (Culioli, 1990; Soga, 2003; de Schepper, 2010). The variety of verbs that can be considered autocausatives is reflected by the high number of occurrences found in the corpora (329 tokens, 26.7%).

The case of anticausative

Similarly to the potential agentless passive, the anticausative diathesis is marked by the promotion of a patient into a subject position, while

suppressing the agent and/or cause of the process. However, there is a semantic nuance between these diatheses coming from the fact that **anticausatives typically dismiss the intervention of an actor with consciousness and volition** (Fellbaum & Zribi-Hertz, 1989). A potential agentless passive sentence such as *Le français se parle ici* hints either the possibility or the obligation for an agent to carry out the process, even if not mentioned; meanwhile, an anticausative such as *Le verre se casse* seems to indicate a process happening spontaneously. There is no causal effect that can be recollected in the sentence, hence the denomination anticausative (Labelle & Doron, 2010).

Historically, the **cause** in anticausative sentences could be expressed as an external argument *par*-phrase, in the same way that reflexive passives licensed expression of the agent (Heidinger & Schäfer, 2008):

(17) *Tout vin donc ne s'aigrist pas par la chaleur.*

(18) *Le contraire se montre par les guerres.*

While it is no longer possible to express causes in this manner in all types of anticausative sentences, some Canadian dialects still allow their use, testifying to the historical use of *par*-phrases with pronominal verbs (Authier & Reed, 1996). Furthermore, many verbs accept causes as external arguments in standard Modern French as adjuncts in prepositional phrases, similarly to conversive passives, turning the clause into a **marked anticausative** (Heidinger, 2015):

(19) *Le fromage s'est moisi rapidement avec l'humidité.*

Typically, anticausatives license inanimate patients undergoing a self-

explanatory process, in which case these patients are also known as themes (Jackendoff, 1992):

(20) *Les os [...] se cassent plus facilement.*

(21) *L'héritage des droits d'auteur se cristallise dans les traités.*

(22) *La seule lueur d'espoir qui se détache actuellement de la situation anarchique [...]*

Other events characterised as anticausative include aspectual verbs (*se terminer*, *se prolonger*, *se maintenir*) and descriptive verbs (*se définir*, *s'avérer*, *se révéler*) (Heidinger & Schäfer, 2008). Anticausatives made up for 16.8% of annotated diatheses in the corpora (208 tokens).

2.2.5. Insights and limitations

The results we obtained from our analysis of the diathetic mappings based on the most common uses of our initial sample of 20 representative pronominal verbs, as well as the corpus annotation of 1,232 structures, enabled us to classify pronominal-verbal constructions under eight potential diatheses. The use of Kulikov's (2011) typology of diatheses helped us find out how each mapping could be interpreted and how these verbs could be compared with other verbs at a more general linguistic level. The view of diathesis as a systematic pattern of arguments enabled us to “dissect” complex structures, in the same way that construction grammar encodes syntactic structures with their semantic layer, argument by argument (Goldberg, 1995). However, we found diathesis to be a more convenient way to build up a classification, since construction grammar is often criticised for its overgeneration of information (Van Trijp, 2015).

The eight diathetic mappings we sketched, while varying in terms of complexity, seem consistent enough to make a case for a potential classification of pronominal-verbal constructions. We believe that a classification based on diathetic structures has the following additional advantages:

-Methodological consistency: It interconnects semantic, syntactic, and valency analyses to deliver consistent argument mappings, in the same fashion as Goldberg's (1995) construction grammar, but with the aim of generalizing patterns into a typology.

-Accounting of lexicalised structures: As a consequence of this method, any verb can have its pronominal form explained, since a diathetic mapping operates at several linguistic levels. Lexicalised (or semi-lexicalised) pronominal verbs (such as *se rendre compte*, *se prendre un coup*), in particular, could be explained on the basis of an empirical analysis when purely semantic or purely syntactic analysis fails to deliver a satisfying explanation (see Gaatone (1975) and Burston (1979) for the rationales behind lexicalised pronominal verbs).

-Accounting of polysemic verbs: In particular, the focus on uses rather than forms allows us to connect several verbs that would otherwise be treated as separate cases (such as *se casser*, usually classified as subjective because of a focused view on *Le verre se casse*, whereas *se casser la jambe* exists along with *se laver le visage* as reflexive benefactives).

Our analysis was also the opportunity to confirm that these verbs, which are seemingly heterogeneous, share a common property: a reduction in valency of the verbal process. Even in the case of the reflexive benefactive diathesis, the original dative meaning is suppressed when using the *se* variant. As

Melis (1990) originally stated, this also means that the reflexive pronoun *se* has no other function than modifying valency, being “*formally an accusative without becoming a grammatical argument*” (p.125). Further insights into this issue can be obtained by comparing the diatheses we found for French pronominals with Kulikov's (2011) definition of the middle voice, which includes a variety of diatheses often accompanied by an intransitivisation of the neutral structure or, in other words, a valency reduction. The researcher writes that the middle voice encompasses **the passive, conversive, anticausative, reflexive, reciprocal, antipassive, and reflexive benefactive** diatheses. It is not fortuitous that these diatheses are, for the most part, represented by the same morphological mark in Russian (reflexive *-sja/-s'*), but also by French pronominal-verbal structures, as shown in the classification we have built up in the present study. This serves as further evidence that pronominal-verbal constructions are characterised by a valency decrease by one degree compared with their neutral diathesis variant. The fact that this valency reduction can apply to different semantic roles – agent, patient or dative – appears in our work as one of the most prominent properties of the pronominal-verbal constructions. Although this observation was already hypothesised by several authors (Dik, 1985; Bień, 2007; Creissels, 2007; Corminboeuf & Heyna, 2015), no attempt had previously been made to confirm this valency reduction based on a systematic analysis.

Comparison of the classifications of the sample verbs

Once we took into account these specificities of our diathetic classification, we finally attempted to compare how the 20 verbs from our sample would be classified under the different classifications mentioned in Section 2.2.2.

Several observations can be made from looking at the comparative table

found in Annex 1. Firstly, reported authors' interpretation of *se boire*, *se vendre* and *s'utiliser* as passives do match into the category of potential agentless passives. Likewise, most canonical reflexive diathetic uses are referred to as reflexive or active verbs in the other classifications. Reciprocal verbs are also consistent with our analysis of canonical reciprocal diathesis.

The controversial subjective category encompasses several different structures for each verb. Our analysis of these forms as either anticausative, autocausative and impersonal passives allows us to make a clearer distinction among the subjective uses. It should be noted that the distinction between the autocausative and the anticausative still requires a good understanding of the semantic layer of the arguments, since these diatheses are identical from a syntactic point of view.

Limitations

Our diathetic classification and its eight categories reveal recurring patterns that are general enough to be used for a quantity of verbs, while being brought together by the valency operator *se*. However, even though this classification has its merits, it still should be tempered by several facts.

Firstly, the degree of complexity of diathetic mappings still implies some degree of interpretation, which means it involves subjectivity. As such, a thorough understanding of notable uses of pronominal verbs is often necessary to determine, for instance, the semantic roles of the arguments.

Secondly, while we paid attention to consider as many pronominal-verbal structures as possible, especially in the corpus annotation task, it is possible that certain pronominal-verbal uses might not perfectly fit one of the

categories we built up in this work. In such a case, it would be interesting to figure out how frequently the said structure is found, and how its diathetic mapping could be constructed to be incorporated into our typology.

2.3. Conclusion

This chapter was the opportunity to describe the theoretical foundations of the experimental project we aim to carry out in this dissertation. **The first half** of the chapter was focused on exploring CBI and its roots, as well as uncovering the effects of the methodology in second language acquisition. We performed a systematic review of CBI classroom experiments, which shed light on the overall soundness of the methodology for L2 concept acquisition. **The second half's** objective was to depict the target concept of our study, namely pronominal verbs in the most thorough manner possible in order to construct our SCOPA. We first looked into previous classifications of pronominal verbs which were made in the domain before designing a new type of classification based on the concept of diathesis. Finally, we carried out an annotation task of two corpora in order to verify the consistency of our work.

By combining the tenets of CBI with the newly built taxonomy of pronominal verbs, we should theoretically be able to make a proper SCOPA in a schematical form. While we opted for another format in the end – as will be described in the next chapter –, here is what a flowchart-type SCOPA would look like for our typology of pronominal verbs (Figure 7):

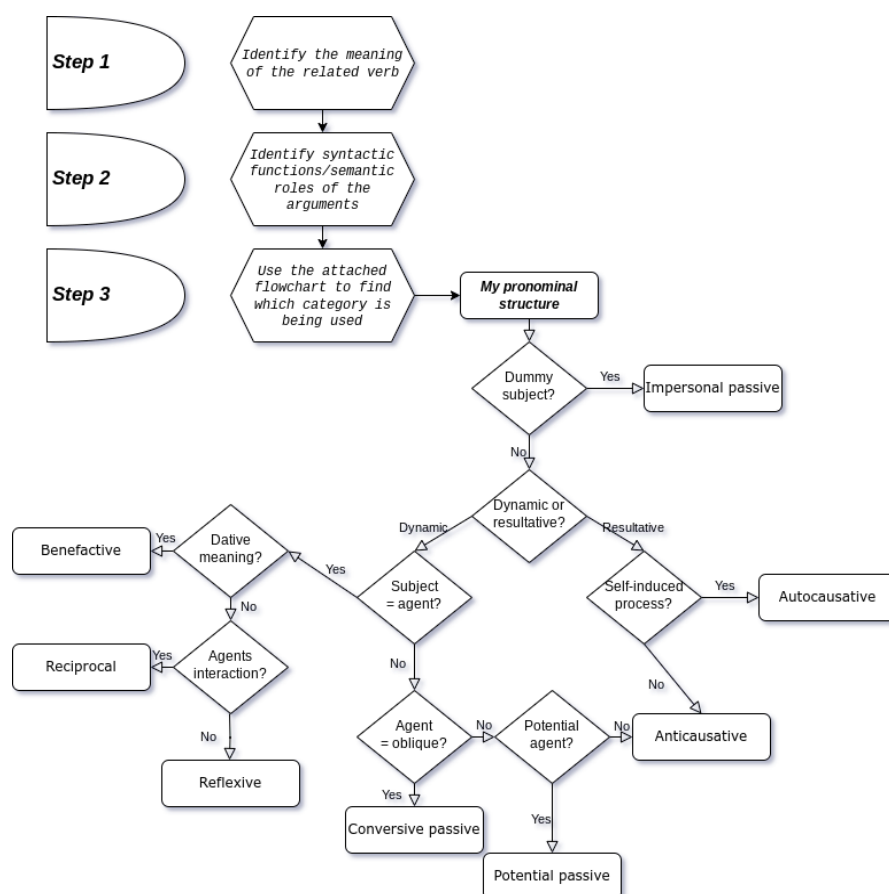


Figure 7: Flowchart representation of our 8-type taxonomy of pronominal verbs

Several shortcomings must be mentioned about this format of SCOPA. First and foremost, this type of flowchart would most likely be insufficient, as many grammatical notions such as semantic roles and syntactic functions are probably unknown to many language learners. This also conditioned the way we had to build our constructs, as we could not integrate all parameters at will for the sake of readability. Finally, the specific context that we had to face when experimenting our module – which will be explained in the next chapter - implied that the use of a flowchart alone was judged insufficient to

provide a satisfying CBI environment.

All in all, as we look to testify the effects of CBI on the acquisition of French pronominal verbs with Japanese university students, this theoretical chapter hopefully offered a solid background on the tenets which we will use as the backbone of our experiments. The following two chapters will focus on describing our methodological choices: first by explaining the systems we developed, including the integration of our typology into a brand-new type of SCOPA, then by relating to the experiments carried out in Japanese universities.

Chapter 3: Research Methodology – System Design

Before addressing the design of the experimental module in Chapter 4, we will aim to explain the way CBI was applied to the typology of pronominal-verbs based on diathesis we constructed. In order to do so, we will first focus on the motives behind our choice of SCOBAs.

As mentioned earlier in this work, a CBI module is usually constructed through the combination of 3 phases: materialisation of the target concept, verbalisation of the materialised concept and internalisation of the verbalised concept. Materialisation phase's main feature is the SCOBAs, or **visual/sensorial representation of the concept** taught through the instruction module. As the SCOBAs should be as thorough as possible, researchers and instructors should have, prior to the experiment, a deep knowledge of the target concept, which is the reason why we went through, in the second half of Chapter 2, a review of classifications of pronominal verbs, the creation of a new classification as well as research on its generalizability through corpora annotations.

Clarifying the concept was merely the first step in the building of our SCOBAs. Indeed, the method used to materialise our classification of pronominal verbs had to be carefully examined. At first, the idea of a traditional classroom exchange in which a graphical representation of pronominal verbs – as it was shown in Section 2.3 - would be explained was looked into. However, the uncertainty in the global context of Covid-19 pandemic outbreak led us to think outside the box in order to carry out our instruction module, as we could not conduct long-term classroom exchanges. Additionally, representing our typology of pronominal-verbs with a flowchart without sacrificing thoroughness would end up requiring a much

too complex diagram, which could hardly be used in a motivating, instructive way for students. By taking inspiration from other CBI works in which visual and/or digital support was used (Serrano-Lopez & Poehner, 2008; Suñer & Roche, 2019), we chose to create a mobile game as our SCOPA.

Mobile gaming or, in a broader sense, digital gaming for language learning (DGLL) has been studied as a potential tool for teaching ever since the number of users in the video game industry started to drastically increase (Hung et al., 2018). Digital natives grew up along the popularisation of modern technology in the 21st century and are comfortable with games and other digital devices. There is evidence that mobile games, in particular, generate affective and cognitive learning outcomes, notably increasing motivation for students to learn (Schmitz et al., 2012). Mobile games boast incredible popularity in the West but also in Japan as in 2019, 73% of Japanese aged 18 to 24 were playing games on their smartphones on a regular basis¹⁴. The practical utility of a digital support in the context of the Covid-19 pandemic outbreak, as well as the familiarity of Japanese learners with mobile games and the overall possibilities with the format were arguments that drove us to consider it as the support for our SCOPA.

Digital games allow for interactivity while giving the player independence and free choice, making them a complex yet interesting choice for language teaching research. This is especially true for language learning, as games are established as prominent grounds for immersion and use of the target language (Su et al., 2021). The prevalence of interactive tasks as powerful in-game features led us to opt for Task-Based Language Teaching (TBLT) as

¹⁴ Statista Research Department: Feb 4, 2022

a comparative design relative to CBI, and to develop a second game inspired by this approach. The first half of this chapter (Section 3.1) will be first focused on covering the possibilities which emerged in the field of DGLL, the state-of-the-art of TBLT in serious gaming as well as the impact of teaching methodologies on the building of games. Notably, we will investigate how CBI and TBLT have been incorporated into games. Meanwhile, in the second half of this chapter (Section 3.2.), we will describe the experimental design linked to our work by presenting the development tool that we used (RPG Maker), the creation and content of our games as well as the evaluation methodology we chose to carry out our experiment.

3.1. Digital gaming for language learning

This section will briefly introduce the most prominent features of DGLL as well as explain how both TBLT and CBI have been incorporated in games.

3.1.1. Learning through gaming: state of the art

The development and the success of the video game industry has brought strong interest for researchers on the effects of gaming in learning. Several developers, such as EduGame, choose to explicitly make games for learning, and a small share of video games are labelled as serious games – games with a strong educational objective as well as thematics referring to the real world. However, the vast majority of video games are still cultural products focusing on the entertainment of the player (Karasimos & Zorbas, 2019). It does not mean that these games could not be used to learn, because video games naturally give opportunities to the player to learn by "*confronting the player to a voluntary interactive activity, in which [they] follow rules that constrain their behavior; enacting an artificial conflict that ends in a quantifiable outcome*" (Salen & Zimmerman, 2003: 80). This interactive

conflict also ensures that the player's cognitive abilities are required in order to finish up the task.

Generally speaking, experimental research on the effects of video games for learning is divided into two categories: research focused on popular games originated from the entertainment industry (*entertainment games*, EG) and research based on a game explicitly created or modified for an educational aim (*games for learning*, GL). In their systematic study of DGLL, Acquah & Katz (2020) mention that out of the 26 papers covered, only 6 were carried out with an EG. Indeed, EG's focus is not on linguistic acquisition, so it can be hard to quantify the immediate effects of these games on language development. Using GL to teach language knowledge seems to be more convenient for researchers as these games deliver a list of tasks adjusted to learners' skills (Sykes, 2013). However, Sykes also cautions that GL might undercut the recreational aspect of gaming in favour of a serious tone, therefore reducing the interest that learners could take in the game, while an EG offers an authentic gaming experience in which language learning sequences are observed from the outside without affecting players' motivation.

Although DGLL games were also experimented in other regions (such as Western Europe, in Cornillie et al., 2012; North America, in Cobb & Horst, 2011), reported DGLL research has been mostly carried out in Eastern Asia (China, Taiwan, Korea, Japan), Northern Europe (Finland, Ireland) and in the Middle East (Iran) (Acquah and Katz, 2020). Additionally, a comparative study done by Su et al. (2021) shows that most studies in DGLL target the acquisition of English, followed by Chinese. The prominence of English as target language could be due to the international nature of the language in the fields of leisure and entertainment.

The genre of games was also a noteworthy factor. 54% of the studies

reported by Acquah & Katz are categorised as educational games, while the other half includes strategy games, simulation games and role playing games. These genres are also the most common in Zou et al. (2021)'s study on digital game-based vocabulary learning (DGVL). The authors also notify that researchers in DGLL avoid using action, fighting or sport games because of negative representations (such as violence) or the lack of linguistic content. However, there does not seem to be a correlation between the game genre and learning outcomes. In contrast, several game features were reported to have a strong impact on the acquisition process. In this section, we will briefly describe the impact of these features on linguistic acquisition, before showing how they apply to the TBLT and CBI games.

Impact of game features on linguistic acquisition

Reinhardt & Thorne (2020) identify game-related L2 learning as a combination of the following features: (1) contextualisation and linguistic environment, (2) time and iterative play, (3) shelter for practice, (4) goal orientation and purpose, (5) languaging and sense saturated coordination, (6) identity performance and cooperation, (7) independence and spatial mobility, and (8) autonomy.

Among the most important features that enhanced vocabulary learning, **autonomy** (a strong degree of player control) was often cited as very impactful (Bibauw, 2022; Butler et al., 2014). Unlike in a regular classroom setting, digital games create a self-contained environment where the target knowledge should theoretically be learnable without additional support (Reinhardt & Thorne, 2020). In practice, L2 gaming requires at times integrated support such as subtitles, access to glosses and dictionaries. Yet, it is still possible for the player to engage and progress through a DGLL game at their own pace, a factor which greatly contributes to motivation and

reduces the stress of learning (Acquah & Katz, 2020). Autonomy was notably important for linguistic acquisition, as a high degree of autonomy was statistically correlated with improvements in vocabulary learning in several works ($d = .326$ in Ebrahimzadeh & Alavi, 2016; $t(49) = -2.728$, $p = 0.009$ in Yang et al., 2016).

Cooperation was also a powerful factor, strongly impacting learners' motivation. A number of studies emphasise the way the video game World of Warcraft (Figure 8) offers an environment favouring multilingual communication as well as player literacy through cooperative settings (Thorne et al., 2012; Zheng et al., 2009). According to Peterson (2010), the degree of inhibition to communicate, and thus, cooperate is lower in an online game than in a classroom setting, which implies that a video game is suitable for creating interactive situations advocated by Task-Based Language Teaching (TBLT). Butler et al. (2014) also claim that interactions made in virtual worlds are an important source of intrinsic motivation for learners, even if this motivation could be tied to the game's content rather than on linguistic development. As such, games are suitable to provide powerful contexts for learner cooperation.



Figure 8: Games like World of Warcraft promote cooperation as a way to carry out tasks.

On the other hand, **competition** was a controversial factor for acquisition outcomes. It can at times "enhance motivation, but it also was found to distract the player from the learning or even demotivate players" (Acquah &

Katz, 2020: 12). Competition could indeed push players towards a goal and build a sense of teamwork, but also garner negative emotions from them, such as distress and anxiety.

The feeling of **challenge** or difficulty was also cited as an essential aspect of DGLL in the 26 papers. Whenever the game confronted learners with a task considered too difficult, they usually felt anxiety, leading to a slow down in the learning. Meanwhile, a task considered too easy would negatively impact their motivation (Acquah & Katz, 2020). This statement allows us to compare gaming tasks difficulty with Vygotsky's ZPD (1987), as a proper balance between player's skills and game difficulty should create an optimal experience in which learners' focus is intensively required while allowing them to visualise their success. Players-learners should feel confident in their ability to carry out the task, especially since games are effective tools for learning offering students a hypothetical environment in which they can explore alternative decisions without the risk of failure. The creation of a space in which decisions can be made with lowered risks and consequences, known as the *psychosocial moratorium*, is especially relevant in language learning because linguistic acquisition is tied with exploration and practice – processes naturally encouraging mistakes and guesses (Gee, 2007; Vogel, 2018).

The game features mentioned hereabove were applied to a variety of games, built with different theoretical foundations. Especially, as digital games are often focused on the notion of accomplishing an objective or a task, they are naturally associated with one of the dominant frameworks for teaching foreign languages, Task-Based Language Teaching (TBLT). The following sub-section briefly details the way TBLT is incorporated in DGLL and presents a series of experimental games which were designed as TBLT.

TBLT applied to digital games

TBLT is characterised by the distinction it makes between tasks and activities (Ellis, 2003). A **task** is an activity requiring language use, with a pedagogical aim and leading to various outcomes. Since a simple drill exercise could be considered as both a task and an activity depending on the context surrounding its outcomes and application, defining the boundaries between a task and an activity is not always easy - especially when designing a teaching module. Furthermore, activities considered as tasks can also either be "real-world" or "pedagogic" (Ellis, 2014). While real-world tasks feature strong communicative aspects by mimicking situations found in real-life (such as taking a reservation at a restaurant, etc.), pedagogic tasks only focus on generating natural language use (but do not necessarily involve communicative situations). Consequently, it is not always clear whether games designed with the intent of offering a task-based experience are authentic TBLT paradigms, nor can we say whether the tasks they include are real-world or pedagogic material.

Vegel (2018) writes that TBLT uses several features common with video games in order to distinguish task-based learning from exercises: rules, goals and objectives, outcome and feedback, conflict, competition, challenge and opposition, interaction, and representation of a story. As such, a digital game can also have some sort of **pragmatic and/or communicative objective** as well as ties to the real world, meanwhile a language learning book's only purpose is pedagogical. According to Schmoll (2016), leaving the possibility for learners to embody a character in a virtual world motivates them further into participating actively in TBLT tasks. Compared to stable supports such as textbooks, games are *"far more flexible at not only affording players a sense of agency but also affording options for difficulty adjustment leading to more individualised learning by design."* (Vegel, 2018: 117). As such, a TBLT sequence should focus on putting learners in situations in which they

can only progress with proper, personal involvement (Salaberry, 2001). Due to the fact that they push the player in different directions depending on their choices, tasks such as **branching dialogues** are the essence of TBLT interactions found in digital worlds and enhance the feeling of control that the player has over his gaming experience (Bibauw et al., 2019). While branching dialogues can take many different forms, they are mostly based on the principle that each play-through experienced by players is different depending on the choices they make, ensuring non-linearity of a game. Other tasks found in TBLT games requiring player involvement include interactions with the environment, quick-time events, or puzzles. A major asset of such in-game tasks is that feedback from any action is instantaneous, allowing the player to see the consequences of any decision they make (Vegel, 2018).

As the ties between DGLL and TBLT could suggest, there were a number of works in TBLT involving the development of a video game. To verify their effects, we were mainly interested in uncovering task criteria (Vegel, 2018) or, in other words, how these works attempted to incorporate linguistic content through in-game tasks, as advocated by TBLT tenets. Indeed, even if DGLL work normally features prominent aspects correlated with linguistic development in gaming, such as impactful contextualisation, balanced difficulty or autonomy in learning, the nature of the tasks leading to such development varies. In *Lost in the Middle Kingdom* (Shepherd et al., 2011), the player is asked to solve puzzles and interact with non-playable characters (NPCs) in order to advance to further stages of the game. *I-FLEG* (Amoia et al., 2011), *Broken Sword* (González et al., 2021) and *Crystallize* (Cheng et al., 2017) are other examples of games where the player has to explore pre-rendered environments and/or inquire NPCs in order to progress. Other

games such as *Food Force* (Hitosugi et al., 2014) and *Daily Verbs* (Aljayyousi, 2021) do not allow for free movement, and are focused on delivering linguistic content through textbox choices. In general, these variations in the degree of player autonomy - often cited as one of the most important features in DGLL (Butler et al., 2014) - did not seem to bring great differences in efficiency. Furthermore, deHaan (2005) suggests that games with a high degree of controllability for the player, when combined with contextualisation and pattern repetition, lead to beneficial linguistic development. However, even though giving more control to learners has its merits, a number of studies also warn that when in-game tasks are too demanding, players tend to have weaker gains than observers of said tasks (deHaan et al., 2010; Amoia et al., 2012; Chen & Yang, 2013).

In-game tasks can also be analysed for their way of depicting linguistic knowledge. In the introductory chapter of this thesis, we mentioned how the way target concepts are conveyed is not something always agreed upon. In the context of DGLL, there does not seem to be a consensus either whether games should include as much exposure to the L2 as possible, or rather propose L1 explanations to warrant learners' comprehension. Indeed, while *Broken Sword* and *Lost in the Middle Kingdom* alternate text written in the target language (L2) and translation content in learners' first language, other games like *Crystallize*, *Food Force* and *Daily Verbs* exclusively use the target language. Even then, these games can sometimes make use of the first language to facilitate linguistic explanations, because sufficient pre-task linguistic knowledge is not always properly balanced between learners. This is the case in *Lost of the Middle Kingdom*, where translations are available for dialogue minigames. A post-experiment survey even showed that some students expressed a desire to see more parts of the game explained in English for ease of understanding, highlighting the fact that TBLT games

might need to bestow sufficient salience in players' L1 to be effective. For that matter, Pinter (2005) and Shintani (2012) suggest that the use of L1 should not be completely banned in TBLT tasks, but rather progressively diminished as learners progress through. Indeed, Shepherd et al. (2011) reported that the major obstacle in developing the story in their game was to avoid using players' L1 as much as possible. Even if that was compensated by heavy use of 3D models, imagery and visual effects, authors admitted that the ban of L1 felt limitative towards developing complex linguistic content. Finally, the length of in-game tasks was the last diverging factor in TBLT games. *Food Force* included 6 missions totalling 1 hour of gameplay (which could be repeated up to 5 times), which resulted in significant improvements in the long term (Hitosugi et al., 2014). Even longer playtime was experimented in *Broken Sword* and *Lost in the Middle Kingdom*, where learners were asked to play a maximum of 5 hours per week for 3 weeks. Both games showed significant linguistic gains after the testing period. Meanwhile, Amoia et al. (2011) experimented a game with shorter sessions (2 sessions of 5 minutes of tutorial and 15 minutes of gameplay) which led to mixed results. Several learners complained that the sessions were not long enough, as they were asked to find out-of-place objects in a house within a limited time. Sufficient time-on-task should be considered for optimal results, as granting additional time for pre-task planning has been shown to benefit accuracy, fluency and complexity through digitally-assisted L2 learning (Ziegler, 2016).

All in all, these features hint at the way TBLT tasks can be incorporated into games, albeit a lot of these tasks are closer to pedagogic tasks rather than real-world tasks (Ellis, 2014). The next step of our research would be to focus on the features found in CBI compared to CLT/TBLT, so that we can determine the differences required in designing the two games which will be

used in our experiments.

3.1.2. Applying CBI to gaming

Since CLT generally puts the focus on learners rather than instructors in the teaching process, role-play activities and technology-assisted learning are commonly used (Dos Santos, 2020). Most CLT-inspired games are based on completion of small tasks and are identified as TBLT games. As mentioned in the previous section, TBLT games can take many forms and often lead to successful improvements in language skills. They involve tasks requiring the use of the target language, such as solving puzzles, interacting with characters or reading letters (Shepherd et al., 2011).

On the other hand, CBI-inspired games are less defined in the literature. Theoretically, they should be focused on conceptual development rather than development of practical, communicative skills. Even then, Fehrmann (2016) argues that TBLT and CBI should be compatible, notably by integrating the materialisation phase in pre-task activities and the verbalisation phase in post-task activities. While the researcher attempted to do so in an experiment with prefield elements in German, it is much more difficult to collect such data in digital games. In Shepherd et al. (2011), the authors make a clear distinction between TBLT-based task accomplishment, required to progress through the game, and CBI-inspired knowledge content, presented as diagrams, images and maps and promoting conceptual acquisition. This mix of CBI elements with other tools was also experimented within a CBI case study on the acquisition of German cases (Pienkoss, 2021). In that case, the SCOPA consisted in the use of a game application akin to *Pictionary* to describe linguistic situations varying according to the German case. While the application was praised for its benefits, it also required additional linguistic explanations taking the form of

classroom exchanges. This highlights the fact that the engineering of a balanced SCOBAs is a difficult task, especially when using a support with non-linguistic content. The systematic review of CBI studies we made (cf. Chapter 2) shows in fact no report of games using SCOBAs in classroom experiments up to 2025. Even though several studies made successful use of multimedia technology to deliver their CBI module (Serrano-Lopez & Poehner, 2008; Fehrmann, 2016; Suñer & Roche, 2019), digital supports such as video games are seldom used to teach grammatical knowledge, even outside of the CBI field (Su et al., 2021).

On the other hand, we did find examples of games focused on lexical knowledge which made use of learners' L1 as a mediation tool, albeit they are not considered as proper CBI works because of the absence of SCOBAs and thorough conceptualisation tasks. This is the case for *Tasukete Gotou!* (Steel, 2020), a game developed for the acquisition of advice-giving vocabulary in Japanese. While the game was entirely built in the target language, verbalisation exercises, in which learners discussed the game's content, allowed them to switch between English and Japanese. Learners were also asked to think about advice-giving sentences in their L1, in order to emphasise the differences with the L2 in later stages of the experiment.

While L1-intensive games are, for obvious reasons, uncommon in the field of second language acquisition, the idea of mediating thoughts with the L1 in order to develop L2 knowledge is, in fact, in line with Vygotsky's principles (1987). This is why we considered making use of L1 mediation for the acquisition of a grammatical concept as complex as pronominal verbs. As mentioned earlier, the present research is the opportunity for us to investigate the effects of CBI on the acquisition of pronominal verbs in comparison with TBLT by using video games. Two games will be developed for this matter: a

CLT/TBLT-inspired game serving as a control tool, and a CBI-inspired game in which this role of the L1 as mediator will be further explored than usual. Synthetically, the two games will have the following differentiating features:

Feature	CBI-inspired game	CLT/TBLT-inspired game
<i>Game language</i>	Mostly L1	Mostly L2
<i>Replayability value</i>	No	Yes (difficulty choice, randomly generated exercises)
<i>Target knowledge depiction</i>	Conceptual (understanding of properties mixing linguistic, audiovisual and gameplay content + verbalisations)	Implicit (through task completion, although a glossary is available for comprehension purposes)
<i>Game structure</i>	Linear	Non linear
<i>Typology used</i>	8-category diathetic typology	Traditional 4-category typology (Zribi-Hertz, 1987)
<i>Objective</i>	Internalisation (ability to use pronominalisation without help of the game)	Task completion (in particular finishing the final task)

Table 19: Overall comparison of our CBI and TBLT games

As can be seen, the difference in the features included in our games was motivated by the constraints of their respective theoretical approaches. To integrate these aspects into our games, we used a software named RPG Maker MZ. The following section will detail its characteristics, its limitations and show how it was used to build the two experimental games.

3.2. The use of RPG Maker MZ as developing tool

The design of the two games highlighted in the present work will be detailed through this section. After presenting the software used to build both games, we will precisely explain their features, but also the design limitations we encountered while making them.

3.2.1. Presentation of RPG Maker

RPG Maker MZ is a software developed by Degica© and released in 2020. It allows the developer to create a role-playing video game by mixing simple scripting language with pre-made built-in resources. RPG Maker MZ uses JavaScript as scripting basis, which encourages acute development and increases the possibilities in coding. The interface of the built-in resources in RPG Maker MZ is divided into two main categories: a database, which allows the developer to assign data such as names and values to game entities, and a mapping section in which it is possible to integrate said entities in a generated environment. The combination of pre-made data and JavaScript-based data grants the software a harmonious touch and empowers a feeling of “easy to use, hard to master”. RPG Maker MZ allows the developer to create environments (interconnected “maps”), characters, dialogues, and generate many actions available to the player.

Several research projects were carried out using RPG Maker in SLA. DeHaan (2011) combined project-based learning and video games with Japanese EFLs. In one activity, students were asked to create their own role-playing game using RPG Maker VX (a previous instalment in the RPG Maker series), which is modeled after the video game *Final Fantasy X*. The students started by providing a summary in English of *Final Fantasy X* and using it as the basis for their own game; they then put together an interactive story and looked up scripts of other games that matched the types of game they wanted to design. This activity gave deHaan a way to talk meaningfully about story genres, English dialects, and dialogue; he also had students design an online English gaming magazine, which afforded them the opportunity to develop reading, writing, listening, and speaking skills as well as technology skills such as digital publishing. In another study, Hwang et al. (2017) developed a problem-based English listening game with RPG Maker,

which was tested with 77 Taiwanese ninth graders (English L2). Gaming materials common in role-playing games, such as battles, exploration rewards, treasure and weapons were associated with learning materials like English listening tasks, dictionary, vocabulary listening cards and instruction advice. The 77 learners were divided into two groups: 39 students were tasked to use the experimental game while the 38 others had to learn the same content while listening to the music player and through the use of slides. The study revealed not only differences in degrees of motivation and the feeling of anxiety between the classes (ANOVA: $F = 26.24$, $p < .001$, $\eta^2 = 0.26$) but also overall higher listening achievement ($F = 17.53$, $p < .001$, $\eta^2 = 0.19$). The correlation also showed that students with higher English anxiety and lower scores in the pre-test had significant improvement which levelled in with the lower English anxiety, more proficient students. As these studies showed satisfying results, RPG Maker appeared to be an appropriate choice to be used in our game design, especially thanks to the previous experience we'd accumulated with this software.

3.2.2. RPG Maker Games design

Two games known as Pronominia were designed for the research: a TBLT-inspired game used for the comparison group, and a CBI-inspired game used for the experimental group. Even though both games were designed as *"serious games for autonomous play, and hence foreseen for out-of-class voluntary use [...]"* (Bibauw, 2022: 80), they were still created with different approaches in mind. This sub-section will detail the features of these games. Before that, we will detail the technical aspect of the deployment of the games for mobile devices, the difficulties we encountered and their resolution.

3.2.2.1. Deployment of the RPG Maker Games

RPG Maker allows developers to deploy their projects into several forms, while encrypting audio and graphic data with a developer-chosen key. The Windows version is the most standard output format and is immediately usable after using RPG Maker's built-in deployment function. By contrast, the Mac OS, web browser, Android and iOS versions are more complicated to operate with, as the built-in function merely generates a folder of encrypted files.

For the pilot experiment (october 2023), we initially chose to develop 4 versions, with 2 target versions being the Android and iOS versions, and the Windows and Mac OS versions working as substitutes (in case of technical problems, or should any student be unwilling/unable to play with their mobile phone). Indeed, the original idea was to offer the possibility to play the game offline on any support, to ensure optimal conditions for learners. However, there were concerns about the collection of save data when using offline versions, but also other technical difficulties tied to differences in devices.

Consequently, for the final experiment, which took place in May 2024, we ended up making a standardised, web-hosted version playable on any device as long as users were connected to the internet. The HTML version ended up working the best, as it would allow students to keep their playthrough automatically saved as long as they kept playing on the same device, while not requiring any download or installation. As a result, pictures found in the following sub-sections were taken out of the HTML versions of our games.

3.2.2.2. CBI-inspired variant

The CBI-inspired game is an adventure game with a few role-playing game (RPG) elements. The player can move around, speak to non-playable characters (NPCs), consult an interface in which an inventory of concepts is available. With the help of a duology of Japanese context/French dialogues, the CBI game follows the story of Tanaka, a Japanese businessman helping out a French detective, Mr. Dupont, in order to solve a case. As the player progresses through the game, he's confronted to “dream worlds” depicting diathetic categories of pronominal verbs as they were detailed in Section 2.2.4.2. For instance, the first “dream world” embodying the canonical reflexive diathesis is called “The Castle of Mirrors”. In this castle, the player witnesses strange phenomena as well as NPCs requiring help. To solve the mystery and escape the castle, the player has to understand canonical reflexive verbs as a concept, use them as interaction tokens with the environment or NPCs, and answer questions based on their meaning or utility. A lot of effort was put on the building of these “dream worlds”, which would each time take the essential characteristics of our diathetic typology of pronominal verbs. According to Negueruela (2008), this step of materialisation of concepts as didactic tools should be thoroughly worked out, as it is essential for the concept to acquire a salient status for learners.

The game is a story-driven adventure revolving around quest-based chapters. Each quest is mostly focused on one category of pronominal verbs related to the typology we constructed, although some chapters include more than one category. Unless the player finishes up a quest, he is unable to progress further, enhancing the fact that this game is a progression-based game (Juul, 2005). In between these quest phases, there is generally an interlude during which Tanaka (the protagonist embodied by the player) studies, goes to work

and walks around the city in search of hints to help Mr. Dupont. Unlike the quest sequences focused on the acquisition of pronominal verbs, in which the player has to move around, inquiry NPCs and think about a way out (cf. Figure 9), these interludes do not generate much interaction for the player other than following the story texts and answering punctual questions about the case and/or conceptual knowledge linked to pronominal verbs.



Figure 9: "Dream world" sequences allow the player to move around and interact with NPCs

Since the player has relative freedom to explore until they find the key to progress through each chapter, such as interacting with NPCs in a certain order or retrieving a concept-item in the environment, this game is effectively a **situated avatar-based game** (Cornillie & Desmet, 2016). This means that it emphasises the development on a holistic level of the four major L2 skills, and that it is less focused on drills and repetition than shorter games known as minigames. In such an environment, a major challenge was to integrate verbalisations naturally without changing the flow of game. As we've seen through the review of literature in Section 2.1., a CBI module cannot be considered as such without giving the opportunity for learners to

verbalise the newly acquired knowledge. To this aim, we chose to use two strategies: the integration of verbalisations into each materialisation sequence, and the deliberate access to a post-game chapter for a more explicit way to verbalise. To be precise, each and every dream world sequence - which are part of the materialisation phase – not only presented each time a new aspect of the target knowledge, but also restrained the players so as to prevent them from moving on in the game without completing specific quests. These quests were an opportunity for learners to **mentally verbalise**, as they required them to “*process the concept with the help of the SCOPA, then without it*” (Renoud, 2016: 66).

An example of that would be the way “The Castle of Mirrors” (representative of the canonical reflexive diathesis) involves the player in discussions with NPCs who feel obligated to watch themselves carry out an action through their reflection in the mirror (see the sequence shown in Figure 10). The player previously received knowledge and guidance about the canonical reflexives from the king of castle – who explained the schematic elements of the **SCOPA corresponding to this category** (in this case: *dynamic action, subject = agent, no dative meaning, no agents interaction*). Notably, the player will gradually obtain **grammatical notions as items** (such as agent, dative, etc.) which can be consulted any time, and even used at specific instances. Then, each learner must understand the action performed by the NPCs and depending on the situation, they might be able to help them solve a problem.



Figure 10: Example of SCOBa realisation

In this sequence, one of the NPCs is shown as falling asleep (s'endormir) as he is verbalising this action while viewing the mirror. The interaction allows the player to obtain a new glossary entry (in Japanese: [s'endormir obtained!]):



In the Castle of Mirrors, all characters are acting in a similar fashion, actualising their verbalised pronominal verb with actions. However, another NPC (black-haired man), is panicking because he forgot what action he was supposedly doing (represented by the broken mirror):



This goes on until the player realises and tells the black-haired man that he actually stopped (s'arrêter) watching himself in the mirror, actualising the canonical reflexive's meaning alongside with his own action!

This innovative way of integrating verbalisations more naturally is made possible with the interactive nature of DGLL, as the game flow only progresses along with actions of the player.

For every world corresponding to a diathetic category, a similar pattern will occur:

- 1) The players will be confronted to category-representative behaviours (NPCs having to interact with each other for the canonical reciprocals, greedy NPCs for the canonical benefactive, NPCs fighting for leadership (agency) in a village for the agentless passive, etc.);
- 2) They will need to find information (taken from the schematic version of the SCOBAs) regarding these behaviours/diathetic categories and understand their characteristics/why they are expressed that way, notably by collecting words and notions;
- 3) Finally, they will find out a way to leave the dream world by repairing any anomalies which will be found – usually directly connected to a specific nuance taken by a pronominal verb (just like the NPC found in the Castle of Mirrors *stopped* watching his reflection because of his action of *s'arrêter*; two NPCs in the world of reciprocals will be acting separately because they argued (*ils se sont disputés*), etc.).

While these activities combined materialisation of the target concept and verbalisation (mostly throughout actions), we also decided to include a more straightforward verbalisation task to complement this approach, as part of a bonus chapter available only to learners who had completed the game as prerequisite. In this chapter, the player has to express their own definition of diathetic categories, as well as the essence of pronominalisation by

reconstructing a puzzle. One way of doing so is to reconstruct the desired meaning by progressively adding limitations to the sentence (Figure 11). In a way, they are constructing their own version of the SCOPA in the process. There is no right answer here; their choices will be registered as a sequence, which will be then used by the researcher as part of data when needed.



Figure 11: Optional verbalisation-type tasks are available once the main story is completed (in Japanese: [Patient] [Verb] ...)

From a broader perspective, the CBI-inspired game follows a three-way process: materialisation of the concept [main content of the game] → verbalisation [post-ending content of the game] → internalisation /acquisition. However, even if the concept is materialised - thus explained - before the learners are able to verbalise, we would still avoid considering the acquisition process tied to this game as *deductive learning* (Vincent et al., 2013), because the border between explanation and task is not easy to apprehend in the case of games (which are tasks by nature, cf. Vogel, 2018). Meanwhile, the internalisation aspect of the game is embodied by the connection between the representation of the target concept and in-game tasks, as they use a *tool-and-result* pedagogy by nature to involve the learner and enhance awareness (Negueruela, 2008).

The playthrough structure of the CBI variant of Pronominia is as follows:

Game part	Tasks	Expected length
Chapter 1: Mr. Tanaka, businessman	Information processing (reading)	25 min
Chapter 2: First step into the dream world – The Castle of Mirrors (<i>canonical reflexive diathesis</i>)	- Information processing (reading) - Use of in-game controls (finding the right path) - Interaction with NPCs (Dialogue acts including Assert, Information-request tasks) - Interaction with the environment (Use of an in-game tool to repair a mirror)	20 min
Interlude 1: Helping out Mr. Dupont	Information processing (reading)	15 min
Chapter 3: Second trip to the dream world - The village of Pairs ; the Mansion of Selfishness (<i>canonical reciprocal diathesis; reflexive benefactive diathesis</i>)	- Information processing (reading) - Use of in-game controls (finding the right path) - Interaction with NPCs (Dialogue acts including Assert, Information-request tasks) - Interaction with the environment (Use of in-game tools, understanding a riddle)	35 min
Interlude 2: Decisive trails	Information processing (reading)	10 min
Chapter 4: Third trip to the dream world – The Forest of Passives (<i>agentless passive diathesis; conversive passive diathesis; impersonal passive diathesis</i>)	- Information processing (reading) - Use of in-game controls (finding the right path) - Interaction with NPCs (Dialogue acts including Assert, Information-request tasks) - Interaction with the environment (Use of in-game tools, understanding a riddle)	40 min
Interlude 3: Is the case finally solved?	Information processing (reading)	10 min
Chapter 5: Last trip to the dream world – The Island of Causes (<i>autocausative diathesis; anticausative diathesis</i>)	- Information processing (reading) - Use of in-game controls (finding the right path) - Interaction with NPCs (Dialogue acts including Assert, Information-request tasks) - Interaction with the environment (Use of in-game tools, understanding a riddle)	30 min
Interlude 4: Uncovering the truth	Information processing (reading)	15 min
Bonus chapter: Have I improved? Testing my new understanding of Pronominal Verbs!	- Information processing (reading) - Verbalisation task (Use of in-game tools, Information-request tasks)	~ min
		Total: 200 min

Table 20: Detailed content of the CBI-inspired game

Game design choices

Typologically, the game includes **conversational tasks** (information processing with some branching dialogues) but also **other types of interactive tasks**, ranging from the interaction with the environment to specific use of in-game tools (see Ellis, 2003 for a definition of conversational tasks). The branching dialogues in Pronominia (cf. Figure 12) are part of a narrative system, which means that there are several preset outcomes with the in-game dialogues (Bibauw et al., 2019).



Figure 12: Branching dialogues also appear in interludes

We chose to use this semi-closed system because it allows relatively easy data processing while granting a decision factor to the player. Every choice is registered in a database accessible for the researcher. As for the language used in the game, we chose to use Japanese for the major part of the dialogues and scenes; however, the learning sequences in the “Dream Worlds” and dialogues with Mr. Dupont include a variety of French sentences. Not only is it in line with CBI's theoretical tenets, which argue

that *linguaging* is best done in learners' L1 (Kambaru, 2023); this was also decided so that the effects of the L1 could be analyzed solely in terms of improvements on conceptualisation, as insufficient lexical knowledge – and thus, miscomprehension - could hamper a process heavily mediated by the understanding in one's L1 (Renoud, 2016). Inside the game, the branching dialogues were either connected to the story (e.g. making a decision about following a trail), or linked to Tanaka's thoughts about pronominal verbs (in which case, the player has to guess which form/meaning is correct in a given situation). Even though these tasks were important to solidify the player's interactions with the game, they would not be as prominent as in the TBLT-inspired game. Indeed, CBI puts the focus on the understanding of a concept through a variety of means, tasks being merely one of them.

For this game, we focused on giving a lot of **autonomy** to the player, as it is one of the most contributive factors to efficient language learning (Acquah & Katz, 2020). It needs proper balancing, as the higher the degree of autonomy is, the more difficult making the right choices might become. This is why we opted for an adventure game with an interactivity level typical of narrative systems (Bibauw et al., 2019). This type of interactivity ensures, through the use of branching dialogues as well as interactions with NPCs or the environment, that there is enough autonomy for the player, even if it is tempered by the fact that the choices are limited by a progression-based system, which is linear by nature.

Game time (time on task) is another important variable which needs to be balanced appropriately. While interludes were short and straightforward, each learning section required substantial time to be completed, as can be seen in Table 20. Some sections (such as chapter 4) had to be longer than

others because they'd cover more complex diathetic categories. In each chapter, the player is free to take up more time and reiterate previous dialogues with NPCs, because self-pacing is essential for optimal L2 learning (Reinhardt & Thorne, 2020). Additionally, the time invested while finishing each individual part grants the player a sense of achievement, an essential factor for motivation (Acquah & Katz, 2020). Other design choices were in line with CBI tenets, such as the use of learners' L1 to facilitate the understanding and the verbalisation of the target concept, the use of learners' target language for production sequences and the use of our diathetic typology of pronominal verbs constructed as a SCOBAs.

In order to collect adequate data of learners' experience in the game, we created a series of variables that were modified accordingly to their playthrough. The collected in-game variables are detailed as such:

Variable	Part of	Measure	Used for
Game time	Time on task	Minutes (min)	Calculating potential correlations between game time and acquisition
Playthrough n°	Playthrough	Number (0+)	Measuring learners' desire to replay the game
Total answers (branching dialogues)	Performance	Number (0 to 15)	Measuring exposure to pronominal verbs
Total correct answers (branching dialogues)	Performance	Number (0 to 15)	Measuring correctness and precision with pronominal verbs usage
Health gauge	Performance	Number (≤ 100)	Informative about learners' overall in-game choices relative to the story
Good ending	Performance	Number (0 to 1)	Informative about learners' overall in-game choices relative to the story
Bonus tasks	Performance	Number (0 to 5)	Informative about learners' tendencies, curiosity, desire for challenge
Verbalisations	Performance	Number (0+)	Measuring the number of times the player attempted to verbalise a pronominal structure in the bonus chapter

Table 21: List of extracted in-game variables (CBI-inspired game)

In this sub-section, we showed how we integrated CBI tenets into a digital game, notably accounting for the importance of conceptual content and verbalisations. Next, we will present the features of the TBLT-inspired game.

3.2.2.3. CLT/TBLT-inspired variant

The CLT/TBLT-inspired game is a simulation game designed with the aim of *"fostering the meaningful, task-based practice of the target language in communicative settings, and thus to contribute to the development of the learners' L2 productive proficiency [...]"* (Bibauw, 2022: 97). Prior to the building of the game, one concern was that traditionally, TBLT does not have the best affinity with grammar teaching, because TBLT content typically is based on communicative tasks (such as starting a conversation with a clerk, making an order, etc.) rather than broad, linguistic objectives such as "learn how to use this structure in French". Indeed, the nature of tasks provides an additional challenge when considering the design of a complex TBLT module (Soyoof, 2018). In particular, digital games require sound and empirically grounded SLA principles to enter TBLT framework, because they offer a much higher degree of autonomy than many other methods (Ziegler, 2016). TBLT tasks are student-centered and work best when providing autonomous learning contexts. In order to have an adequate control setting for our experiment, we'd have to construct tasks which could be tied with concrete objectives promoting interaction while portraying the grammatical concept in a satisfactory fashion. As a matter of fact, we considered the option of using another method derived from CLT, namely task-supported language teaching (TSLT). However, TSLT is necessarily associated with a pedagogical phase which serves as an introduction to the target knowledge (Ellis, 2014). Consequently, the task is often not self-sufficient for teaching a complex concept; it's only used as a support rather than a stand-alone activity. Due to our wish to make a self-sufficient game including all the required information, we believed that TBLT was a more appropriate design.

As Vogel (2018) reminded, digital games can hardly be considered as simple

learning activities, because they are necessarily interactive, unlike books. Feedback obtained from within the game is immediate and allows the player to change their approach in a dynamic manner at any point. Furthermore, games are designed as cognitive tasks in which input-based exploration and output-based execution are mixed in, while keeping these aspects as implicit as possible (Philippette, 2014). While playing games is by nature an activity outside of the classroom putting the attention on pragmatic meaning (Willis, 1996), such as advocated by TBLT, in-game tasks are also potent in delivering **reasoning-gap** activities, in which learners are able to decide for the best course of action in a given situation. These assets are best embodied by one specific in-game task we created for the TBLT game, called **final task**, which requires the player to test their understanding of pronominal verbs *implicitly* while in reality, they are playing a challenging turn-based game with strategic decision-making and reaction timings (Figure 13).



Figure 13: The final task is modulable and can be accessed from the park

In addition to smaller tasks linked to the theoretical content which also provided instantaneous corrective feedback, the completion of this final task would effectively be an appropriate TBLT content, due to the fact that completing a game requiring a strategic mind as well as attention is a powerful motivator and a daily task by itself. As a consequence of this design, linguistic acquisition would not only be an objective of the game sequence but also a means to achieve other goals (such as finishing the story segment, increasing one's performances in-game or mixing up one's playthrough). Theodorsen (2015) wrote that using language as a means to accomplish an in-game goal is an ideal representation of a task. Accordingly, this task and the others included in the game would fit Skehan's (1996) typology for a TBLT approach, albeit as pedagogic tasks rather than real-world tasks, while the game itself represents a real-world task (Ellis, 2014). To be precise, as was shown with other TBLT games we described in 3.1.1., elements which can be considered as tasks rather than activities are fairly common in DGLL, because these gaming experiences naturally push players in different directions through personal involvement (Salaberry, 2001). These elements were integrated into the three traditional phases of a TBLT module (Ellis, 2000):

TBLT phase	Game asset
Pre-task	Introduction to the game world with the narrative
Task	Learners play the game by completing the proposed sections (mini-games)
Post-task	Game offers feedback on learners' performance and encourages replayability

Table 22: Integration of game assets into TBLT phases (Ellis, 2000)

Instead of unfolding a deep storyline similarly to the CBI-inspired game, the TBLT-based game is focused on delivering a light narrative context instead.

For this narrative, we chose to enhance the light-hearted impression given by the *comics* aesthetics of RPG Maker base library assets. At the start of the game, the player is invited to embody two children discussing and looking for a play in a playground. The child performing the best at the play they chose would enjoy an ice cream offered by the other. Since the narrative setting does not require deep levels of understanding, it ensured that the player had immediate access to task contents, but also helped in justifying the integration of challenges and rhythm games – which children are more likely to play.

To increase one's chances at the mini-games, the player is advised to enhance the comprehension of pronominal verbs. The children will then be able to take lessons about verbs and other matters by speaking with various characters in the theater before attempting to play the game they've chosen. These lessons can be accessed from a list of sections phrased as types of pronominal verbs (reflexive verbs, reciprocal verbs, passive meaning verbs and subjective verbs). Indeed, in contrast with CBI's necessarily more complex diathetic typology of pronominal verbs due to the fact that a SCOPA has to explain the target concept thoroughly, the categories used in this game to explain the variety of pronominal verbs were chosen based on Zribi-Hertz's work (1987), which was presented in Section 2.2.. This made more sense as a diathetic category could not be implemented without massive use of explanations in L1, as well as requiring a dynamic, conceptual in-game representation, which would be counter-intuitive for a TBLT module focused on tasks and direct output/input. To be precise, Zribi-Hertz's version of the traditional 4-type typology was not only justified by its acknowledgment in the domain; it was also selected for the clarity of the explanations as well as its simplicity (Bień, 2007). At first, the learner can only access theoretical learning, taking the form of interactions with NPCs

as well as reading sections with humoristic phrases and sketches (cf. Figure 14a, Figure 14b); however, after completing a theory segment, it is possible to carry out its respective quiz mini-game from the main menu (cf. Figure 15).



Figure 14a, Figure 14b: Optional theory lessons in the theater mix in visual representations with practice-oriented rules of thumb



Figure 15: Quizzes are optional activities used to practise the understanding of pronominalisation

For the most part, quizzes require the player to understand whether a verb highlighted in a translated French sentence should be in a pronominal form. As corrective feedback is an essential asset of TBLT due to the fact that learners should always be aware of their decision-making (Ellis, 2003), every answer selected by the player receives immediate feedback, announced by a buzzer sound (in case of wrong answer) and a positive sound (in case of correct answer). At some point, the character might randomly become tired, and a real-time mini-game (commonly used in gaming industry and known as QTE, *quick-time event*) will appear. Should the player hit the correct timing, a motivation scale will increase and the player will be able to perform better in the final task. If the mini-game is failed, the player character will lose some motivation and the final task will become a little bit harder each time. This system allows the player to stay engaged in the

quizzes, rewarding concentration and inviting learners to increase their character's motivation as much as possible to complete the final task more easily (Halbhuber et al. 2023). After discussing with NPCs to learn about the theory behind pronominal verbs and practicing, the children should be able to go through to the final task confidently, and show off their knowledge to their peer in the playground. The task has 3 degrees of difficulty, each represented by a symbolic play as a "foe": novice (a trampoline), intermediate (a swing), advanced (a pond, to play pebble bounces). The difficulty was initially selected by the player in the introductory part of the game. In a similar fashion to rhythm games for language learning such as *RoboLingo*®, the final task is a mini-game requiring the player to react to verbal content showing up in real-time on the screen by choosing an adequate action between *Pronominal* (if they believe it's a pronominal verbal sentence), *Non-pronominal* (if it's not a pronominal verbal sentence), *Wait* (up to 3 times) and *Use an item* (if they possess one), inside the RPG Maker integrated battle scene. Each action has an impact on the flow of the minigame, leading the player to lose energy points, recovering them, or progressing through the scene by taking a foe's energy points (Figure 16). Should the player's energy points reach 0 beforehand, they will be forced to withdraw and a game-over screen will appear. To make things easier, the player is encouraged to complete quiz sections, as they will not only influence the motivation scale, potentially increasing the character's total energy points, but also reward the player with recovery items (which can be used to increase one's chances of success). In case of failure, the player will be invited to start the game again, and to go through other sections and/or reduce the difficulty level. With the help of items and wait points, the task is designed to be dynamic and can be handled with flexibility; this flexibility participates in the structuration of a compelling TBLT experience, because

player's decisions have an impact on the outcomes of the game.



Figure 16: Final task battle scene

The higher the difficulty, the longer the sequence becomes and the less time the player is given to process each verbal content. This also means that the player is encouraged to play the game again if they wish to complete all three degrees of difficulty. Even in the lowest difficulty mode, the task was designed to be hard to complete without having a good understanding of pronominal verbs (as testified by Japanese learners of French who tested the game in its development phase), which is why the player is highly encouraged to go through the other sections preably. Additionally, completing a learning section grants the player an item to recover energy points, further bolstering the interest for the learner to complete the sections before trying to tackle the final task. Using these items or not is up to the player's choice, granting them the option to challenge themselves if they wish to. The differences in difficulty are the following:

<i>Difficulty degree</i>	Novice	Intermediate	Advanced
<i>Maximum number of verbs</i>	100	110	120
<i>Target hit points</i>	30	35	40
<i>Verbal content show-up time (seconds)</i>	4	3	2

Table 23: Difficulty degree differences (final task)

Thanks to the sequential cut of the game between categories, scenes leading to the final task, theory sections and/or quizzes can be freely accessed from the main menu. As such, most of the story content is directly connected to the final task. The game is thus not a real progression-based game, because the player is able to complete the final task and end the game at any point if they feel confident in doing so (Bibauw et al., 2019).

Game part	Tasks	Expected completion time (min)
Introduction	Information processing (reading)	5 min
Les verbes pronominaux réfléchis	Information processing (reading) Branching dialogues, quick-time events (quiz)	5 min >10 min
Les verbes pronominaux réciproques	Information processing (reading) Branching dialogues, quick-time events (quiz)	5 min >10 min
Les verbes pronominaux à sens passif	Information processing (reading) Branching dialogues, quick-time events (quiz)	5 min >10 min
Les verbes pronominaux subjectifs	Information processing (reading) Branching dialogues, quick-time events (quiz)	5 min >10 min
Final task: I can do it!	Information processing (reading) Rhythm mini-game, branching dialogues	25 min
		Total estimated time: >60 min

Table 24: Detailed content of the TBLT-inspired game

Game design choices

We deliberately designed the TBLT-inspired game to be less demanding than the CBI-inspired game. Indeed, the sessions of play are shorter and emphasise repeatability/replayability thanks to a mini-game driven design (Cornilie & Desmet, 2016), rather than heavy narrative involvement or conceptual understanding, not unlike other TBLT-inspired games such as *Food Force* (Hitosugi et al., 2014) and *Daily Verbs* (Aljayyousi, 2021). Games with lesser cognitive demands have the advantage of being more accessible, at the cost of depth in explanations (deHaan, 2005). Since TBLT seeks to promote productive knowledge through completion of small tasks, we needed to design this game as a straightforward, practice-focused experience. Nonetheless, as a TBLT-inspired game, this project had to be adjusted in accordance with the definition of task we detailed earlier (Skehan, 1996). As such, the theory part as well as the exercises had to focus on the **pragmatic uses and understanding of pronominal verbs** rather than lexical knowledge and/or precision. In particular, exercises had to be constructed in such a way that the explanations in the theory part would suffice to hint at the meaning of each sentence. A summary of each theory content would become available in a glossary after completion and accessible at all times so that students can feel comfortable with each task (Figure 17). The **accessibility of explanations when needed**, provided that the player unlocked the respective sections beforehand, blurs the otherwise common *inductive* paradigm of TBLT, in which learners are required to construct their own rules by practice and task completion (Vincent & Lefrançois, 2013). Yet, the concept is never materialised in an exhaustive way akin to the CBI-inspired game, so that learners typically need to handle all required tasks numerous times before they might feel that they understand the nature of pronominalisation.

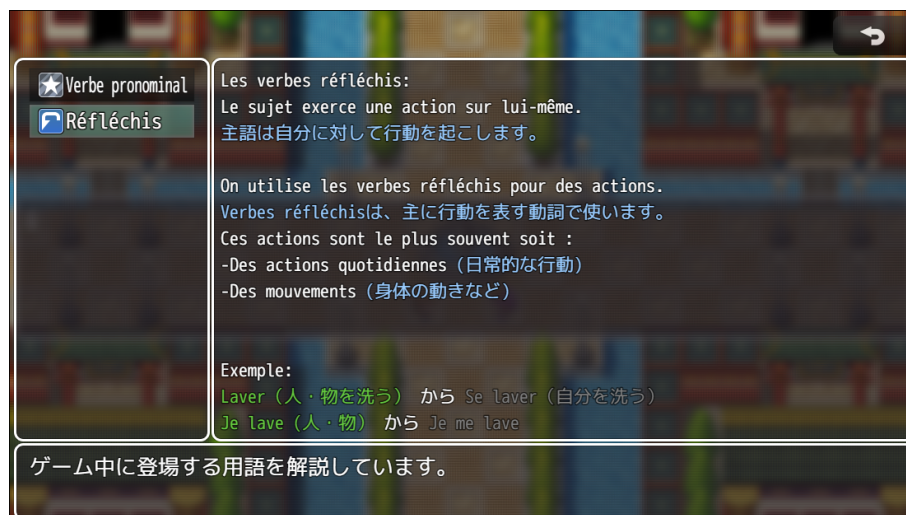


Figure 17: glossary entries fill as the player progresses through optional sections

Another important tenet of TBLT comes down from the fact that **interactions need to be engraved into reality**. This should not only apply to the main interaction, which is the game itself as a real-world object, but also to the interactive tasks included in the game by making sure that the material provided is not fallacious. Quizzes were randomly generated for each part from lists of 100 examples for each type of pronominal verb, taken from the corpora *Sequoia* and *FTB* that we annotated in earlier stages of this research (Section 2.2.). For each category, we manually took 100 sentences out of our 1,232 occurrences without strict category matching; e.g., exercises found in reflexive verbs quiz did not necessarily only include occurrences accounted in the canonical reflexive diathesis, albeit they were a majority. In cases where sentences were too long to be put in a text box, we simplified them by removing unnecessary complements. Inside the game, a random number would then be attributed to each sentence, generated at each attempt at the quiz to offer variety, with the only limitation being that the same

number could not be shuffled twice in a row. The use of corpora-extracted sentences ensured that the quiz sentences were likely based on real-life material rather than constructed sources. Yet, to make sure that the task would be accessible for learners, we had to apply two principles to the construction of the exercises: offering a **translation for each sentence**, and **limiting the CEFRL proficiency level** of sentences to B1 (as predicted by DMesure; cf. Yancey et al., 2021). Furthermore, mistakes were not detrimental in the pacing of the game, as you could simply go on even after making the wrong choice in a quiz. These mistakes would be registered in the game files while being mentioned to the player through a buzzer sound effect, offering instantaneous feedback.

Task type	Details
Theory sections (Interactions with NPCs, information processing)	Visual and textual information based on simple explanations from Zribi-Hertz (1987)
Quiz sections (Branching dialogues + quick-time events)	Score-type quiz based on corpora-extracted 100 sentences for each category (with translation, CEFRL $\leq$ B1)
Final task (Interacting with the environment; reacting to real-time content; testing one's strategy)	Mini-game with various degrees of difficulty (interacting with verbal content in rhythm, choosing adequate options and managing resources)

Table 25: In-game task types in the TBLT game

The most prominent feature of our TBLT-inspired game, **replayability**, was also the focus of earlier corpus linguistics-based gaming experiences (Chamberlain et al., 2008). Using a corpus of nearly 700,000 English newspaper sentences, the authors built *Phrase Detectives*, a game in which players are required to annotate and validate grammatical objects by moving a cursor. Correct decision-making was motivated by two metrics: a comparative scoring (showing points differential with the Gold Standard)

and a collaborative scoring system (comparing with other users' scores). While the gameplay controls were simple, the huge amount of data covered allowed the game to have a strong replayability. This is also the case in our game, as it incorporates data taken out from *FTB* and *Sequoia* corpora and does not go through a long introductory section like the CBI game. The short and repeatable sessions emphasised in this game mostly factor into two features: **time** and **iterative play** - as they allow the player to have free control over the time spent in-game - as well as creating a shelter for practice, due to the fact that learners are able to practise by themselves at will (Reinhardt & Thorne, 2020).

Finally, we had to find a way to balance language **exposure time** to the TBLT with the CBI game, as it is one of the most important predictors of learning (van Gog, 2013). Specifically, in order to keep track of the time spent by learners as well as the overall number of answers they gave, we chose to create an in-game clock and specific variables, so that we could access this information easily. The collected in-game variables are detailed as such:

Variable	Part of	Measure	Used for
Game time	Time on task	Minutes (min)	Calculating potential correlations between game time and acquisition factors
Playthrough n°	Playthrough	Number (0+)	Measuring learners' desire to replay the game
Total answers (quiz) [1 version for each category]	Performance	Number (0+)	Measuring exposure to pronominal verbs
Total correct answers (quiz) [2 versions for each category]	Performance	Number (0+)	Measuring correctness and precision with pronominal verbs usage; assessing detailed analysis of category-based differences
Final task choice	Performance	Number (1 to 3)	Informative about learners' tendencies, desire for challenge
Final task turn n°	Performance	Number (0+)	Measuring correctness and precision with pronominal verbs usage

Table 26: List of extracted in-game variables (TBLT-inspired game)

We have shown in this section that we attempted to create two games based on sound theoretical foundations. Yet, as much as we tried to balance the two games' design to be relevant for the objective we pursue in this work, they are not without shortcomings. Each game had its fair share of limitations, which will be detailed in the next section.

3.2.3. Limits in game design choices

CBI-inspired game

The most important limitation of the CBI-inspired game comes from the fact that as a heavily story-based game, it suffers from the **narrative paradox**. This paradox is described by Louchart & Aylett (2003) as the fact that the stronger the focus is put on strengthening the story and the characters, the less autonomy the players are given in their playthrough. Since the player has to follow numerous story segments in which the only interaction possible is to read text at one's pace, the gaming experience is not always dynamic. It also means that should the player not follow the story and skip several dialogues, they might end up lost and not know what to do next.

Despite the presence of this paradox, we believed that a story-driven game was the best choice for a CBI design, because creating conceptual environments mimicking reflexive verbs, for instance, would not be possible without offering a proper credible setting. We've seen that conceptual acquisition is usually associated with theories of cognitive embodiment (Borghi & Cimatti, 2010). This implies that activating connections in the neural structure through as many senses as possible could help conceptualisation, notably by mixing visual, auditory and linguistic representations. This also led us to create an adventure in which each diathetic category could be represented through various ways, including meaningful story dialogues, inspiring music and creative worldbuilding.

Another limiting consequence of a story-first, **linear design** is that the knowledge needed to progress is increasingly higher as the story unfolds. As such, the player is expected to be able to understand each diathetic category when attempting to enter the final chapter. It is not possible to go back to a

previous chapter, albeit some information about each diathetic category is still available to consult within the game interface.

Finally, we should also notify the fact that the CBI-inspired game contains about 85% of text written in Japanese, which is **not our first language**. While we used the help of a native speaker and corrective tools to verify all the text content, the sheer amount of textual content in this game (over 10.000 lines of dialogue are included) makes it difficult to keep track of all mistakes. We believe that the overall experience, however, is largely comprehensible and will not pose major problems, as feedback from the pilot experiment demonstrated (Section 4.3).

CLT/TBLT-inspired game

As it was already mentioned earlier, the main limitation we had when designing the CLT/TBLT-inspired game was that the tasks tied to the module are **closer to language tasks than communicative tasks** (Ellis, 2003). Since the foremost objective of the whole sequence was to acquire conceptual knowledge rather than communicative skills, using a methodology usually attached to communicative language teaching principles might not have been the best choice. In that regard, we considered using task-supported language teaching (TSLT) in place of TBLT, due to the fact that TSLT allows prior explicit instruction (Li et al., 2016). However, due to the distribution of participants we wished to include in the experiment, it would be difficult to implement pre-task explicit instructions before giving access to the game. As we'd seen in Section 3.1., several TBLT-based games also focus on pedagogical tasks, enforcing the fact that digital games are, by nature, real-world interactions. Furthermore, grammar teaching **can** be achieved with TBLT by encouraging corrective feedback during task performance (Ellis,

2014), as is the case in the game's quizzes section and in the final task. For all these reasons, we believe that the TBLT approach was relevant as a comparison treatment for CBI.

Additionally, since TBLT is associated with interactive tasks rather than conceptualisation, there were constraints on several dimensions, namely **explicit meaning** and **language exposure**. Indeed, the tasks had to be primarily focused on delivering practical linguistic/non-linguistic content which could be interacted with, and avoid theoretical explanations akin to CBI. Still, we realised that it was not possible to completely exclude the L1 (Japanese) from dialogues, especially because lexical knowledge **should not hamper the development of the acquisition** of the target concept. This difficult balance between practicality and clarity can sometimes be underwhelming for learners, which is why we were very aware of the feedback after the pilot experiment and included several adaptations, such as a glossary with translated explanations.

Limitations affecting both games

As we aim to compare the effects of **two different methodologies** for the teaching of pronominal verbs, the games should only be different in accordance with the theoretical tenets they are based upon. However, one concern might arise from the fact that the typology of pronominal verbs itself was not the same, as we chose to use our diathetic typology for the CBI experiment, while the TBLT experiment was focused on Zribi-Hertz's (1987) more condensed, albeit authoritative typology. In fact, using a diathetic typology as heavy as the one we constructed was not possible, due to the fact that it simply includes too much information and would require additional instruction to the TBLT sequence, notably due to the fact that we tried to

avoid using learners' L1 as much as possible. Furthermore, since we also wished to test typology efficiency as a module-dependent feature, this limitation should rather be seen as correlated with the results, which needed to be evaluated appropriately.

Finally, a major concern came from the fact that both games were distributed with the desire to give as much autonomy as possible to learners. In other words, the mobile support allows them to play the game at their desired pace, even in non-classroom settings. Consequently, **proper game completion could be influenced by other learners**. While it was difficult to prevent this bias in an autonomous instruction setting, it should be reminded that pre- and post-tests were completed individually.

Although they were concerning, these limitations did not prevent us from carrying out the experiments successfully. The next chapter will cover the explanation of the experimental design, including the description of testing methods, data collection process and the elected statistical methodology.

Chapter 4: Research Methodology – Experimental Design

This chapter will unveil the methodology behind our experimental design, which is to say how we attempted to evaluate learning gains tied to the experiments. After describing the participants and the procedure, we will show how the tests were constructed, mainly based on Purpura's (2013) typology of grammar assessment. We will also explain the way two surveys were built, including a vocabulary size test, before addressing the question of the statistical testing methods.

4.1. Participants

The research was first conducted in 3 Japanese universities for a pilot experiment in 2023, then in 5 Japanese universities in 2024 for the final experiment. The population chosen for the study was composed of **students who attended French classes in tertiary education**. The only other prerequisite was that the students **should have had prior basic knowledge of pronominal verbs** (so that they could at least understand their morpho-syntactic construction). Not all the students were major in French; in fact, only the students from Fukuoka University were French majors, while other students had French as optional course/second foreign language. As the selected population was rather heterogenous, it was not possible to precisely control their linguistic proficiency level – which is also a reason why we chose to provide sufficient linguistic support in both games. Age and school year (class) were not controlled either in the selection process, although we did ask for this information in a survey. Naturally, as learners were engaged in tertiary education, we expected learners to be older than 18. In total, 59 learners (pilot experiment) then 109 learners (final experiment) participated in the experiment.

In each class, the students were attributed either game in a **quasi-experimental** fashion. Indeed, we first asked the teachers information about each learner's evaluated linguistic proficiency. Then, we created two groups with relatively even proficiency, randomly selecting through learners of a similar level. Even when considering the quasi-experimental nature of the selection process, it should be noted that the importance put on in-game investment (which will be described in this chapter) diminishes the potential shortcomings of unbalanced samples.

4.2. Evaluation procedure

As mentioned before, in order to measure linguistic development, we chose to adopt a between-group effect comparison through a pre-test/post-test design. According to Marsden and Torgerson (2012), the use of pre- and post-test experimental design can lead to several biases, despite its popularity. Indeed, learners naturally **mature** with time, performing better as they progress in an educational setting; this bias becomes stronger as the time difference between the pre- and the post-test grows bigger (Marsden & Torgerson, 2012). Another concern is called **test effects**; they imply that learners might improve simply on basis of questions raising awareness and triggering learning after the pre-test, independently of the intervention. Finally, a statistical phenomenon called **regression to the mean (RTM)**, first mentioned by Galton (1886), show how extreme measures in a pre-test (in the lower or the upper quartile) tend to move closer to the mean than the results in the close range when the post-test is carried out, as they have a higher standard error on average (Marsden & Torgerson, 2012). As such, learners who performed poorly in the pre-test might improve more than others simply due to chance.

Whenever trying to test the effect of an intervention, researchers should always be aware of these biases and attempt to control them as much as possible. In the present work, we made the following decisions regarding our experimental design to address these biases¹⁵:

- Use of **comparison groups** to ensure internal validity (Marsden & Torgerson, 2012);
- Accounting of **pre-test as a co-variate** in statistical models, so that the statistical power of this variable can be controlled (Kim & Willson, 2010);
- Statistical analyses based on the **effect of variables on post-test scores** to avoid over-interpretations based on the effects of the intervention only (Marsden & Torgerson, 2012).

Pre- and post-tests were separated by the intervention – in the case of our research, a 3-week period for autonomous play. The following figure gives a very broad idea of how the overall experiment was planned out:

¹⁵ A full explanation of the methodological choices behind our statistical models can be found in Section 4.6.

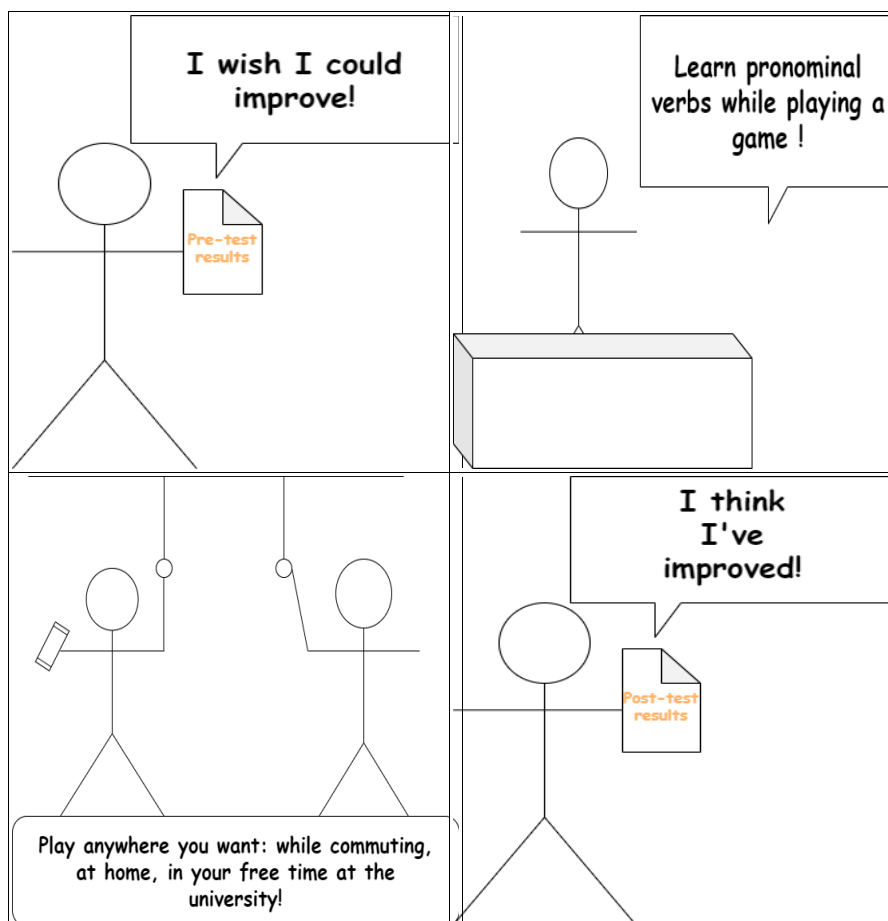


Figure 18: Experimental process

As shown in Figure 18, students were randomly chosen to play one of the two games and were able to progress at their own pace in their free time or while commuting. The implementation of very frequent autosaves (at least at every map change) allows players to keep their progression intact at any point, should they have little time to play. Game saves were also sent to a server so that we could keep track of all learners' game progression throughout the experiment. The collection of in-game variables was complemented with a pre-post evaluation described above and carried out

through online forms. The next section will detail the theoretical tenets behind the creation of the pre-test and the post-test, as well as the other tools which were used to collect experimental data. Notably, we included surveys to measure learners' attitudes and perceptions towards French as well as educational games, as well as a vocabulary test to portray learners' proficiency level in French.

Pre-post evaluations are the main measure of the acquisition of our target concept, which is to say pronominalisation. However, in order to create an adequate evaluation allowing us to measure accurately the development of this conceptual knowledge among students, we first need to elucidate what abilities are attached to it.

4.3. Evaluation material

The acquisition of grammatical concepts entails a broad range of linguistic abilities, ranging from the understanding of the lexical forms and meanings to the development of syntactic awareness, processing and parsing (Jeon & Yamashita, 2014). Pronominal verbs also have specific morphosyntactic properties such as their impact on word order or subject-verb agreement, which is not the case for all types of grammatical concepts. However, as mentioned in the introduction, this study's foremost objective is to assess conceptual understanding - as the ability to judge whether a pronominal verb should be used, as well as factoring the impact pronominalisation has on the sentence's meaning - rather than these morphosyntactic properties, albeit they cannot be ignored entirely. According to Purpura (2013), grammatical knowledge is necessarily a combination of knowledge of form and semantic meaning. This means that suppressing morphology and syntax from the acquisition process could contribute to an incomplete acquisition. Linguistic

knowledge is also affected by the context (ranging from low- to high-context situations) and the discourse; however, the present study will mostly focus on the conceptualisation of pronominal verbs in a low context, sentence level, because we believe that the understanding of pronominalisation is primarily tied to lexical and morphosyntactic changes, as it was shown in our definitions of pronominal verbs in Chapter 2.

Consequently, in the light of Purpura's taxonomy of linguistic knowledge (2013), conceptual knowledge of pronominal verbs would mainly include the acquisition of *lexical forms* (formation/compounding of pronominal verbs), *morphosyntactic forms* (flexion, tense, impact of pronominalisation on word order), *lexical meanings* (polysemy, denotations, formulaic expressions) and *morphosyntactic meanings* (effects of the pronominalisation on the sentence's meaning), as relevant *abilities* measured by activities involving **receptive knowledge, limited productive knowledge and extended productive knowledge** (cf. Figure 19).

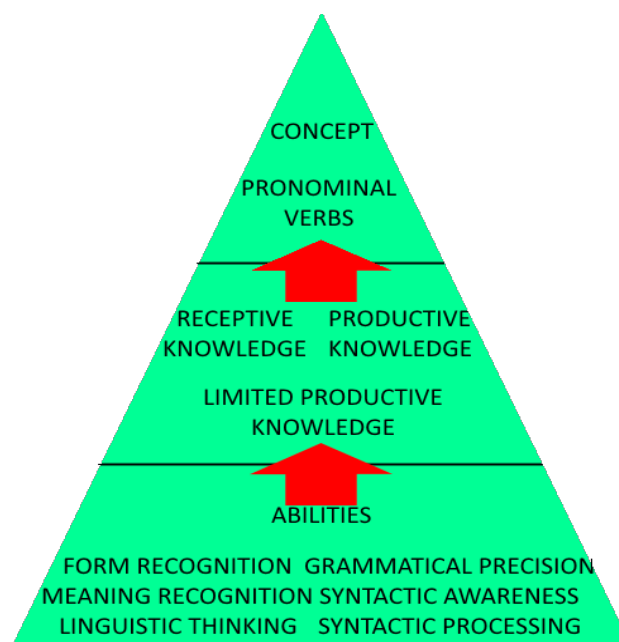


Figure 19: Relationship between abilities, skills and concept

While assessment tasks should focus equally on the acquisition of these four abilities, it should be noted that "*most L2 testers continue to conceptualise grammatical knowledge uniquely in terms of forms, with little or no explicit attention to the measurement of meaning*" (Purpura, 2014: 12). Designing assessment tasks while balancing the weight of forms and meanings is a challenge that we wished to tackle by choosing the following skills, as they can be generalised to both our CBI-based and our TBLT-based experimental modules (Pérez-Tattam et al., 2013; Purpura, 2013):

◆ Receptive knowledge

Receptive knowledge of a given target concept is enhanced by any input collected by the learners throughout the acquisition process. Typically, vocabulary and grammar learning activities are predominantly focused on reception, and the acquisition of complex conceptual knowledge will also

require the use of input-based activities (Webb, 2005). Receptive knowledge would be appropriate to assess the acquisition of lexical and morphosyntactic meanings because, for these tasks, we could create selected answers differentiated by their respective meaning, as well as their impact on the meaning of the sentence without necessarily requiring specific comprehension of forms.

◆ **Productive knowledge**

Productive knowledge of the same concept is enhanced by any output made throughout the acquisition process. In the light of Purpura's theory of grammar assessment (2013), productive knowledge usually encompasses **limited** productive knowledge (in which learners' production is formatted by the evaluation process) and **extended** productive knowledge (in which learners' production is not formatted). Because productive knowledge involves the output made by the learners, related tasks will also put the understanding of lexical and morphosyntactic forms into test (such as the formation of the verbs, subject-verb agreement, adequate tense choice, etc.), even if we will not consider these abilities equally as important as lexical and morphosyntactical meanings when discussing learners' potential improvements in conceptual knowledge.

◆ **Task difficulty assessment**

One major problem with pre-post test designs using different items comes from the difficulty of balancing the tests to have a similar degree of difficulty (Rimmer, 2006). Researchers often choose to have a high number of items while restricting the construct, so that the type of test questions should mostly be homogenised (Buck, 2001). However, this method only covers the basic prospect of assessment, as it obviates the homogenisation

of exercise difficulty.

A solution could be either to take exercises based on textbooks of a same referenced level, or to use automated exercises with standardised settings. Indeed, a substantial number of researchers in NLP (Natural Language Processing) have worked on this issue to provide automation of exercises based on several metrics, such as word familiarity and frequency (Beinborn et al., 2014). For grammar exercises, NLP methods have especially developed automated multiple choice questions, fill in the blank and close exercise items, of which answers are always strongly constrained and predictable (Perez-Beltrachini et al., 2013). Constructing automated exercises with a standardised difficulty measure, such as CEFRL, would allow researchers to create tests with similar levels in an idealistic way. Unfortunately, we were not able to find any validated tool to construct grammar exercises for a specific CEFRL level. Consequently, to design our pre- and post-tests, we chose to generate exercises on the basis of examples found in FSL textbooks of referenced A1 level (*Communication progressive du français, Alter Ego 1, Le Nouveau Taxi*), then adapted sentences to include pronominalisation-related tasks and to be on par with a CEFRL A1 to A2 level, as predicted by the readability tool Dmesure (Yancey et al., 2021). The reasoning behind this choice was that the games were created with a required basic A1 level, in which learners are supposed to have learnt basic uses of pronominal verbs and their formation (Beacco, 2005), without necessarily having conceptualised these verbs in their metalanguage. The complete data of online forms with the predicted average CEFRL level for each exercise can be found in Annex 3 and Annex 4, respectively for the pre-test and the post-test. In the following section, we will describe how we selected certain types of activities linked with receptive knowledge and (limited) productive knowledge in order to build the pre- and post-tests.

4.3.1. Pre-/post-test metrics and design

All the tasks used in the pre- and post-tests were created on the basis of Purpura (2013)'s taxonomy of ways of eliciting grammatical performance; in that list, tasks are regrouped according to their target use (receptive, limited productive and extended productive knowledge). It should be reminded that pre- and post-tests items but different, due to the fact a 3-week instruction period was too short to prevent information retention between the tests.

RECEPTIVE KNOWLEDGE

Selected response tasks (SR) allow us to make inferences about the learners' receptive knowledge of the current learning point. Accordingly, we use SR to check on meaning identification related to the (non-) pronominal nature of the verb (Purpura, 2014). Mackey & Goo (2007) refer to these tasks as prompted responses. We balanced the tasks so that pronominal and non pronominal sentences were equally distributed.

SR1 = *Judgment/Recognition task*

The first SR task requires the learner to choose which of the non-pronominal and the pronominal variant should be used in a series of sentences. The translation in Japanese (which we made with the help of a native speaker) ensures that the correct form will be used. This is a simple exercise assessing the understanding of lexical meanings (denotations of pronominal verbs) and morphosyntactic meanings (impact of pronominalisation – as a morphosyntactic change - on the sentence's meaning). In this series, 5 sentences should take the pronominal form and 5 sentences should be non-pronominal.

Example

Elle va tous les jours à la gare à vélo. La gare [trouve/se trouve] près de

sa maison.

彼女は毎日自転車で駅に行きます。駅は家の近くにあります。

Scoring	Tested parameter
Out of 10 [Correct assessment: 1 each]	Receptive knowledge (Lexical meanings + Morphosyntactic meanings)

SR 2 = *Error detection*

This second SR task is focused on the logic behind the build up of pronominal verbs. As the learners have access to the meaning of the non-pronominal variant in their mother tongue, they should be able to retrieve the meaning of the pronominal counterpart and determine whether such meaning is possible. This is another type of exercise to assess the understanding of lexical meanings (denotations of pronominal verbs) and morphosyntactic meanings (impact of pronominalisation on the sentence's meaning). In this series, 5 sentences are correct and 5 sentences are incorrect.

Example

Ce vin se boit avec du poisson.

「boire : 飲む」

このワインは魚と飲むべきです。

Scoring	Tested parameter
Out of 10 [Correct assessment: 1 each]	Receptive knowledge (Lexical meanings + Morphosyntactic meanings)

LIMITED (CONTROLLED) PRODUCTIVE KNOWLEDGE

Then, we want to verify whether uses of the target concept can be produced in a limited manner (LP, limited production). These tasks require the student

not only to recognise lexical and morphological meanings as it was asked in SR tasks; they also put their understanding of lexical and morphological forms to the test, albeit in a limited and controlled manner (since there is a correct answer for each part of the test). Mackey & Goo (2007) refer to these tasks as closed-ended production tasks, while Norris & Ortega (2000) call them constrained constructed responses.

LP = *Labelling/sentence completion*

This first LP task is focused on form production. Learners are required to identify the correct meaning, similarly to an error detection task; however, they also need to fill the gap with the adequate form (label). It means that precision is an additional factor which has to be taken into account. This type of exercise is useful to assess all four abilities, from the understanding of lexical meanings and forms (denotations and formation of pronominal verbs) to the comprehension of morphosyntactic meanings and forms (impact of pronominalisation on the sentence's meaning, adequate tense, subject-verb agreement, word order). There are 8 pronominal and 7 non-pronominal sentences.

Example

Adapter/S'adapter	Je facilement aux changements de température. 気温の変化に適用するのが慣れている。
Adapter/S'adapter	Le bus la climatisation selon la température. 気温によってバスはクーラーの設定を適用させる。

Scoring	Tested parameter
Out of 15 [Correct variant: 7.5 (0.5 each)] [Correct form: 7.5 (0.5 each)]	Limited productive knowledge (Lexical forms + Morphological forms + Lexical meanings + Morphological meanings)

LP = *Gap-filling/Discourse completion*

Gap-fillers are commonly used in grammar acquisition tasks and allow learners to produce meaningful answers within a controlled environment. Here, the verbs should be chosen from the list below, and then either be used in their pronominal form or their non-pronominal form. However, not all the verbs are supposed to be written down in gaps. This means that learners will have to take into account the context surrounding the gaps to fill. Like the previous task, this type of exercise allows us to assess the four abilities. There should be 5 pronominal and 5 non-pronominal forms to fill the gaps.

Example

「(s') écrire ~ 書く」 「(se) retrouver ~ 取り戻す、合流する」 「(s') aimer ~ ○が好き」

「(se) voyager ~ 旅行する」

[...] Nous[1]..... à l'aéroport avec tous nos bagages vers 8 heures du matin. [...]

[1]

Scoring	Tested parameter
Out of 10 [Correct verb: 0.5 for each] [Correct form: 0.5 for each]	Limited productive knowledge (Lexical forms + Morphological forms + Lexical meanings + Morphological meanings)

EXTENDED RECEPTIVE AND PRODUCTIVE KNOWLEDGE

Finally, extended production (EP) tasks allow us to make inferences about learners' full receptive knowledge, their full production or their overall L2 performance (Purpura, 2014). Unlike LP tasks, these tests give a certain degree of freedom to answer, which also means that there isn't any one-dimensional correct answer.

EP = *Verbalisation*

Example

Expliquez pourquoi un verbe pronominal est utilisé dans la phrase suivante:

Le chien de ma voisine s'est fait vacciner contre la rage.

お隣さんの犬に狂犬病のワクチンを打ってもらった。

This exercise will test receptive and productive knowledge. The freedom learners are given in their answers grants flexibility in the analysis. Depending on the game they've played (CBI-based or TBLT-based), their answer might be different. Whatever the game they've played, they should be able to deliver a convincing analysis. We would expect learners' answers to be coherent and plausible for the task to be successful. More than any other task, this will test their understanding of the lexical/morphosyntactic forms and meanings.

Scoring	Tested parameter
Out of 5 [Convincing analysis: up to 1 for each]	Receptive knowledge, extensive productive knowledge (Lexical forms + Morphological forms + Lexical meanings + Morphological meanings)

While the results of the pre- and post-tests were the most insightful forms of evaluation for the acquisition of pronominal verbs, we also wished to collect data about learners' profile, attitudes and vocabulary size, so as to uncover potential co-variables which could impact the learning process. The next section will focus on presenting the design choices for these other metrics.

4.3.2. Learner profile questionnaire and vocabulary size test

Learner profile questionnaire

Research on learning motivation through DGLL, in particular, shows that learners are more motivated to engage in learning activities in gaming environments (Yang et al., 2017). Due to the fact that some gaming elements are more prone to enhance intrinsic motivation, such as sense of challenge, feeling of fun and immediate feedback, indicators of these features should be gathered as additional co-variables (Iaremenko, 2017). We also chose to collect data about learners' preferences over a specific type of game (heavily narrative, less narrative) as a factor in order to observe latent diverging factors, because a story that is easier to understand has been shown to have positive effects on learning motivation (Chen, 2019; Göbel et al., 2009).

Consequently, in addition to questions about learners' gender, age and university background, we added two series of 15 and 12 questions respectively with 7-points Likert scales to survey their subjective experience prior and posterior to the experiment. Likert scales are indeed an adequate tool to quantify qualitative attributes such as attitudes and perceptions (Joshi et al., 2015). In the first series, carried out prior to the experiment, we tried to extract motivational data relative to their attitudes and perceptions towards French, the target concept (pronominal verbs) as well as games in general. While other theoretical models of motivation in education had their merits for our study, such as the exploration of intrinsic motivation with Ryan & Deci (2000) as well as research on perception of success and failure (Graham & Taylor, 2016), Wigfield & Eccles (2000)'s Expectancy-Value Theory (EVT) appealed to us due to its focus on the perceptions associated with tasks – largely prominent under the form of digital games. To be precise, EVT assumes that achievement motivation and performance are connected to learners' previous experience, perception, beliefs, and

expectations of the value attached to the task at hand. In this regard, digital games have been shown to be correlated with sustained improvements in learning motivation because they enhance the value of said task (Yang, 2012). This motivated us to build up a series of questions relative to these variables. EVT-based questionnaire items should be representative of *motivation scales* (Knekta & Eklöf, 2015), which means that they have to inform about learners' previous **experience**, their **interest**, their **expectations** and the **importance** attributed to the task ahead. Apart from expectations, which were included in the second survey as part of the post-experiment feedback, all these stakes were measured with the first survey.

We also aimed to take away differences related to the game type. While game type is not necessarily correlated with acquisition outcomes (Acquah & Katz, 2020), we still thought analysing the preferences of students could be a useful insight for further research. They will consequently be treated as data relative to task interest.

All these variables will be analysed as independent variables in correlation with our main dependent variables, namely receptive knowledge (all exercises), limited productive knowledge (E3, E4) and extended receptive and productive knowledge (E5).

Variables: [1] Sex, [2] Age, [3] University, [4] School Year, [5] Experience, [6] Interest, [7] Importance

Used as: co-variables for acquisition results

Format: [1] Multiple choice, [2] [4] Number, [3] Nominal, [5] [6] [7] 7-point Likert scale

Length: [1] 1 MC item, [2] [4] 2 Number item, [3] 1 Categorical item, [5] [6] [7] 15 Likert items

In the second survey, administered after the experiment, we designed 12 items to measure perceived **usefulness** and perceived **ease of use** of the

games. Perceived usefulness, also known as extrinsic motivation (Tiwari et al., 2007), is the perception that learners have of the utility of the task and allows us, as such, to verify whether they felt our games were purposeful. Meanwhile, perceived ease of use is a measure of how effort-free the task was carried out. A task with lower effort requirements will generally yield less drop-out risks and enhance the level of intention of use, but might also grant diminished benefits (Tahar et al., 2020). Consequently, observing how users perceived this parameter is also an essential factor to analyze their involvement in the games-tasks.

This series of items was partially inspired by a similar questionnaire made by Bibauw (2022) in a post-test survey about the use of corrective feedback in game-based CALL models. These variables are themselves described by Davis, Bagozzi & Warshaw (1989) as essential measurement scales for user acceptance. This means that learners with positive perceptions of a system's usefulness and ease of use, in particular, should be more prone to accept such a system, to the point that it might affect their performance positively. The parallel of these co-variates with our main dependent variables will attempt to verify any potential correlations. The survey questions were mostly designed around the *Technology acceptance model*, which includes 6 prototypical statements for both perceived usefulness and perceived ease of use (Davis et al., 1989).

Finally, this second survey also included a space for free commentary regarding the game and experiments, so as to gather feedback from students.

Variables: [1] Perceived usefulness, [2] Perceived ease of use, [3] Free commentary

Used as: [1] [2] co-variables, [3] feedback

Format: [1] [2] 7-point Likert scale, [3] Open answer

Length: [1] [2] 12 Likert items, [3] Categorical item

To ensure the validity of the survey items, we carried out tests of Cronbach's alpha (Tavakol & Dennik, 2011). Cronbach's alpha measures the overall coherence between different items by examining the range of learners' responses, with alphas at 0 meaning that the items are completely unrelated, while alphas at 1.00 meaning that their outcome is perfectly similar. The results of these statistical tests will be shown in Section 5.1, alongside with other measurements.

Vocabulary size test

Along with the first survey related to learners' profile, a small vocabulary size test ($n = 15$) was administered. Vocabulary size has been shown to provide a reliable estimate of L2 proficiency, as the four language skills are tightly connected to vocabulary knowledge, from the perspective of both the output and the input (De Wilde et al., 2020; Milton & Treffers-Daler, 2013; Noreillie et al., 2018). As we lacked information regarding learners' initial proficiency level in French, using a vocabulary size test in the first survey appeared to be an efficient way of measuring said level.

The test was designed as a productive *picture-based* vocabulary size test. It has been shown that receptive vocabulary size is usually bigger than productive vocabulary size (Fan, 2000; Webb, 2008). With the limited number of items included in this survey, we opted for a productive vocabulary size test. Additionally, the use of a productive picture-based test

would effectively remove the effects of guessing responses, increasing its overall reliability (Gyllstad et al., 2015).

For each item, a picture was presented to be translated with precise instructions in Japanese, to be sure that learners would understand which word is targeted. Indeed, pictorial annotations alone may provide too much information compared to precise translations, increasing the noise in the answers (Sweller & Chandler, 1994). Meanwhile, the combination of a pictorial annotation with written content has been shown to have a significant effect on performance in vocabulary tests (Jones, 2004; Plass et al., 1998).

An example of pictorial vocabulary size test item is as follows (Figure 20):



*Figure 20: an example of vocabulary size test item with open answer
(The question is: How do you call this fruit in French?)*

The idea of targeting the word with its overarching category (e.g. raspberry < fruit) decreases the effect of *richness of images* which may otherwise affect learners' ability to accurately transcribe L2 words (Jones, 2004).

The items were selected based on the CEFRL, as reported in Beacco et al. reference manuals (2004) with the help of FLELex's analysis tool for C1/C2

words (François et al., 2014 ; Pintard & François, 2020). 3 items were picked for each CEFRL level, except for C1/C2 (which had 3 in total). The metric that will be used for analysis is the total score, as scores should be lower with less proficient learners (Bibauw et al., 2019). To avoid cultural biases as much as possible, we picked the items from picture provider websites in Japanese and made sure that they had a corresponding existence in Japan. Still, as some pictures could be described by several possible words, we took care to choose pictures referring to entities as precisely as possible (Annex 5 shows the list of pictures which were chosen).

The tests and surveys we depicted in this section and the previous one were the main source of data we used for our experiments. The following section summarizes this data and divides it into the variables which were integrated into statistical models explained in Section 4.6.

4.3.3. Data collection

Including the tests and surveys we described above, the data used in this experiment, except for ID and Treatment distribution, was extracted from five sources:

- a) Pre-test data
- b) Survey 1 data (including vocabulary size test)
- c) Saved game data
- d) Post-test data
- e) Survey 2 data (including free commentary)

All the variables from a), b), d) and e) were first put into an .xls sheet automatically generated with Forms, before we merged them into a single sheet. Key data from the saved game files was then added to the .xls file after being extracted with a command script.

After pondering the use of certain variables such as **in-game scores** and the **number of correct answers made in branching dialogues**, which were described previously as collected data, we decided to leave them out because they would prove to be **difficult to integrate into statistical models** without involving, even partial, subjectivity. Consequently, we ended up with a table including the following 33 variables for each learner:

Variable	Type	Extracted from	Measure	Contains	Used for
ID	Character	-	-	(Student ID)	Indexing
Treatment	Factor	-	-	CBI, TBLT No treatment	All research questions
Sex	Factor	Survey 1	-	Male, Female, Other	All research questions
Age	Factor	Survey 1	y. old	18~	Research question 3
University	Factor	Survey 1	-	(University name)	All research questions
Class	Factor	Survey 1	-	1~	All research questions
Vocabulary Size	Numeric	Survey 1	points	0~15	Research questions 1, 2
Pre-test score E1	Numeric	Pre-test	points	0~10	Research question 2
Pre-test score E2	Numeric	Pre-test	points	0~10	Research question 2
Pre-test score E3	Numeric	Pre-test	points	0~15	Research question 2
Pre-test score E4	Numeric	Pre-test	points	0~10	Research question 2
Pre-test score E5	Numeric	Pre-test	points	0~5	Research question 2
Pre-test total score	Numeric	Pre-test	points	0~50	Research questions 1, 2
Experience	Numeric	Survey 1	Likert scale	1~7	Research question 3
Experience with pronominal verbs	Numeric	Survey 1	Likert scale	1~7	Research question 3
Experience with mobile games	Numeric	Survey 1	Likert scale	1~7	Research question 3
Importance	Numeric	Survey 1	Likert scale	1~7	Research question 3
Importance of pronominal verbs	Numeric	Survey 1	Likert scale	1~7	Research question 3
Importance of mobile games	Numeric	Survey 1	Likert scale	1~7	Research question 3
Interest	Numeric	Survey 1	Likert scale	1~7	Research question 3
Interest in pronominal verbs	Numeric	Survey 1	Likert scale	1~7	Research question 3
Interest in mobile games	Numeric	Survey 1	Likert scale	1~7	Research question 3
Days on task	Numeric	Save data	days	0~	Research questions 1, 2
Time on task	Numeric	Save data	minutes	0~	Research questions 1, 2
Completion	Numeric	Save data	percentage	0~100	Research questions 1, 2
Post-test score E1	Numeric	Post-test	points	0~10	Research question 2

Post-test score E2	Numeric	Post-test	points	0~10	Research question 2
Post-test score E3	Numeric	Post-test	points	0~15	Research question 2
Post-test score E4	Numeric	Post-test	points	0~10	Research question 2
Post-test score E5	Numeric	Post-test	points	0~5	Research question 2
Post-test total score	Numeric	Post-test	points	0~50	Research questions 1, 2
Ease of use	Numeric	Survey 2	Likert scale	1~7	Research question 3
Usefulness	Numeric	Survey 2	Likert scale	1~7	Research question 3

Table 27: List of extracted variables

As we can see, each variable could be used for at least one of our research questions. These research questions will be verbalised in the Section 4.5. Before that, the next section will cover the way our evaluation was applied and the concrete implementation of our experiments in Japanese universities.

4.4. Procedure

4.4.1. Pilot experiment (2023)

Before the practical, larger-scale experimental research we had planned for 2024, we carried out a pilot experiment in late 2023 in order to test out the viability of our module on all fronts. 3 classes were chosen in 3 Japanese universities (total $n = 59$), which were then divided between CBI and TBLT (randomly, albeit with minimal control in order to minimise performance gaps based on individual differences). Every student was asked to complete the pre-test and the preliminary survey, before they were given instructions about the installation of the attributed game.

The experiment period was initially designed to account for a 3-weeks span

in which students should play the game as much as they could, and ideally complete it. However, as mentioned earlier, technical difficulties arose during this pilot phase, as many students were unable to carry out the installation on their device. This not only slowed down the process, leading to ever-growing difficulties in communication with both students and teachers (to the point we eventually had a crucial lack of complete post-test and second survey data), but also drove us to develop another version based on HTML code to enable all students to play the game. While several students mentioned playing the game with a friend by their side in the survey answers, we had no way of measuring it; as such, the participation rate was only calculated to be around 34% ($n = 20$).

Despite these problems, we were encouraged by the fact that we measured improvement between pre- and post-test (cf. Table 28), specifically for the few students who **completed both tests and their assigned game** (respectively $n = 5$ and $n = 3$), as well as the overall interest and positive feedback we have received.

	<i>CBI ($n = 5$)</i>	<i>TBLT ($n = 3$)</i>
<i>Pre-test score mean (/50)</i>	21.90	20.66
<i>Pre-test score mean CI ($n = 50$)</i>	[17.06, 26.74]	[16.50, 24.82]
<i>Post-test score mean (/50)</i>	24.70	26.83
<i>Post-test score mean CI ($n = 50$)</i>	[18.57, 30.83]	[23.49, 30.17]

Table 28: Comparison of pre- and post-test score changes in the pilot experiment (2023)

As we can see, both games would show positive results on score changes, with an especially bigger effect for the TBLT users. However, as the sample was **simply too small** to construct a satisfying statistical analysis, we were

still worried about the outcome while tackling the large-scale research. Nonetheless, due to time constraints and thesis schedule, we chose to carry out our experimental research during Spring 2024.

4.4.2. Final experiment (2024)

Our two different treatments (CBI and TBLT) were meant to use several common assets, such as an interactive gaming experience, accessible cartoony graphics, an autosave function ideal for short play sessions and a choice-based experience. However, as suspected at the start of our research, the difference in conditions also led to diverging results in terms of linguistic knowledge development. This section will focus on detailing the implementation of the experiment which took place in May~June 2024. After presenting the context of the study and the target sample, we will showcase the conduct of classroom interventions as well as data collection.

As mentioned earlier, we previously carried out a pilot experiment in 2023 which resulted in insufficient results due to technical difficulties and lack of post-test data (albeit there was an overall improvement in performance across the board on the little data we had). For the second experiment, we immediately promoted the HTML-based games, as they were easier to use and did not require an installation on learners' devices. We also adjusted a few elements in the game to account for learners feedback, such as the integration of a glossary for the TBLT game.

In late May 2024, we traveled to 3 universities in Japan (Seinan, Fukuoka, Tokyo) and carried out 2 online experiments in 2 other universities, for a total of 5 institutions. The following table details the number of students who participated in the study:

University	Initial number of participants	Status	Experiment launch	Game data collection	Experiment closure	Final number of participants
Seinan University	51	Offline	May 29th	June 19th	June 26th	40
Keioh University	15	Online	May 29th	June 19th	June 19th	14
Fukuoka University	22	Offline	May 30th	June 20th	June 27th	20
Tokyo University	10	Offline	June 3rd	June 24th	June 24th	8
Ryukyu University	12	Online	June 11th	July 2nd	July 2nd	12

Table 29: Overview of the conduct of the final experiment (2024)

The population chosen for the study was composed of **students from Japanese universities who attended French classes**. The only other prerequisite was that the **students should have had prior basic knowledge of pronominal verbs** (so that they could at least understand their morphosyntactic construction). Not all the students were French majors; in fact, only the Fukuoka University students were French majors, while other students had French as an optional course/second foreign language. Age and school year (class) were not controlled for in the selection process, although we did ask for this information in the survey.

As can be seen, there was a disparity in sample numbers depending of the universities. The overall challenge of contacting universities, then travelling to Japan to conduct experiments limited us in the choice of the samples. Also, due to scheduling and budget constraints, we had to organise two of the five experiments in online sessions.

Another limitation came from the fact that we could not use the full data from the initially planned 109 learners; indeed, we'd only received full pre-

and post-test data from 94 learners due to cases of non-attendance in the classes in which the post-test was distributed. Among the 15 learners who did not complete the post-test, there were even learners with substantial game time who will not be accounted for, which was unfortunate for our statistical analyses.

The initial plan was to carry out a single class intervention at the start of the experiment, which included the following steps:

- a) Self-introduction, explanation of the objectives of the research (5 minutes)
- b) Pre-test and first survey (55 minutes) (cf. Figure 21)
- c) Presentation of the games and technical explanations (10 minutes)
- d) Distribution of game URLs



Figure 21: QR codes were scanned by learners to pass the pre-test and the first survey

While the participants were doing the test in step b), we divided the class into 2 randomly composed groups (with the help of the teacher). No particular problem was reported in this phase. A 3-week period was given to learners to play their attributed game as much as possible.

Unlike with the pilot experiment, most learners were able to play the games, and we managed to gather substantial data from their playthrough. Yet, as much as we tried to motivate learners to play, **16 learners who were attributed either game did not even launch it**. As there was no record of in-game variables for these learners, they could not be integrated into the statistical models in a smooth fashion. As such, their test and survey data were used for a third condition, which we named the No Treatment group (NT) and essentially served as a control group relative to our CBI and TBLT treatments.

As the last phase of the experiment, we required all students to pass the post-test as well as the second survey. Thankfully, we were able to rely on the teachers's cooperation to organise this last phase, as we were not able to travel back to Japan as well as due to time zone differences. This last phase, however, was organised at different times, depending on the teachers' availability (as it was shown in Table 29). Furthermore, unlike what was planned initially, we were able to participate in three of these interventions, thanks to the classes being held at later hours. In these second interventions, the teachers invited the students to pass the post-test and the second survey with another series of QR codes, before closing the experiment.

Although the final experiment was carried out more successfully, some aspects of it could have been improved, such as ensuring that all learners would pass both pre- and post-tests, or enforcing students to play the game they were assigned to. However, as we considered autonomy to be an essential asset of DGLL games, it would have been counter-productive to force learners to do something that they were not motivated doing in the first place. As such, we ended up with a population of 38 CBI learners (with

game data recorded), 40 TBLT learners and 16 NT learners (who had 0 exposure to either treatment).

In the following section, we will present the central research questions of this work.

4.5. Research questions

The present study's main objective is to compare the effects of **two different approaches for the teaching of French pronominal verbs**, each with their singular characteristics. To do so, we attempted to identify the processes at work in the acquisition of a concept as complex as pronominal verbs, and envision whether a L1-heavy, conceptual methodology and a L2-focused, mostly implicit methodology lead to different outcomes, with hopes to make them work together in an ideal instruction setting.

As a common ground for the two approaches, we have designed two digital video games (mobile games), with the ambition that this support could garner a strong motivational response. While both games feature an interactive setting with tasks allowing learners to make use of their conceptual understanding of pronominal verbs to progress further, they vary in terms of autonomy given to the player, replayability value, exposure time and concept embodiment. These variations were crafted into each game according to the respective theoretical tenets of CBI and TBLT.

The present research will address the following research questions:

RQ1 Statistical evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, in correlation with game investment metrics

→ The first and foremost research question of the present work reflects upon the impact of game investment on the overall post-test scores, treatment by treatment. The literature has shown several instances of between-group experiments involving CBI in which the experimental group outperformed other groups in terms of gains in conceptual knowledge (Fehrmann, 2016; Lai, 2012; Serrano-Lopez & Poehner, 2008). We hypothesise that CBI-inspired game users should demonstrate, proportionally speaking, better improvements compared to TBLT-based game users regarding the understanding of pronominalisation, although it would ultimately depend on a key factor: **game investment**. In this regard, game time should have the biggest impact on the results, as we have seen that insufficient time on task can lead to mixed results when using DGLL methodologies (Amoia et al., 2011; Ziegler, 2016). Meanwhile, due to differences in design, CBI users should have a longer exposure in a single save, while TBLT users should have a lot of variation in their playthrough data, especially in terms of completion, as the game is non-linear. We would expect both groups to improve nonetheless.

RQ2 Statistical evaluation of each group's post-test performances for each specific exercise regarding the use and understanding of pronominalisation before and after the experiment, in correlation with game investment metrics

→ Just like RQ1, the second research question ponders the impact of game investment on learners performances, but applies these criteria to **each exercise instead of the overall post-test score**. As such, essential differences might arise from this analysis, due to the fact that two exercises are solely focused on receptive knowledge, while the others also incorporate

an evaluation of productive knowledge. We hypothesise CBI users will have significantly better performance regarding extended productive knowledge (E5, think-aloud analysis), as the CBI game includes verbalisation tasks in the game. We could also hypothesise a stronger development of the more difficult pronominal structures among CBI users, as the typology based on diatheses is more precise than TBLT's typology based on Zribi-Hertz's 4 categories (1987). Meanwhile, the TBLT-game included tasks not unlike E1 and E2 of our pre-post-tests, so there might be more noticeable improvements in these exercises for the TBLT treatment.

RQ3 Statistical evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, in correlation with learners survey metrics

→ Unlike with RQ1 and RQ2, RQ3 will focus on the impact of **other metrics which were not directly affected by the treatment**, such as learners' biographical data (age, sex), learners' preliminary attitudes and perceptions towards games and pronominal verbs (experience, importance, interest), as well as the retrospective feedback variables (perceived usefulness and ease of use). We would imagine that learners experienced with pronominal verbs should have an overall better performance. Likewise, learners with stronger motivational indexes (importance/interest) might spend more time with the game and perform better than others. Meanwhile, perceived ease of use and usefulness are well-referenced as factors prominent for a more engaging game experience and potentially, better performance (Nuryakin et al., 2023).

RQ4 Qualitative evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, from a typological perspective

→ Finally, RQ4 will treat the question of the impact of **each treatment's typology** on learners' responses to the post-test compared to the pre-test. Specifically, we will be interested in checking whether each treatment was more beneficial for distinctive categories of pronominal verbs, but also see how each treatment impacted learners' verbalisation tasks (E5). Unlike RQ1-2-3, this question will mostly be studied from a **qualitative standpoint**, which means that we will attempt to see whether differences appear directly within learners' productions. Before the analyses, we would hypothesise that CBI-game users should be more familiar with difficult pronominal structures (such as anticausatives, autocausatives, impersonal passives), as they are discriminated from each other, unlike in the TBLT game (in which they all belong to the subjective category). Furthermore, we believe that the verbalisations (E5) made by CBI users will not use the same type of information as those made by TBLT users, because the CBI-inspired game includes concepts such as semantic roles and syntactic functions which are not present in the TBLT-based game.

As mentioned hereabove, prior to analyzing the results of the experiment, we expect **both games to be beneficial** for the learners in terms of linguistic acquisition, specifically for the understanding of pronominal verbs. While the typology used in the games is different, the aim of both designs was to at hint the situations in which you should and should not use a pronominal verb, as well as the impact that pronominalisation has on a verb's non-pronominal variant meaning; consequently, a lack of effect for any of the

games would imply that we failed to reach our objectives.

All these hypotheses will, hopefully, find an answer after we carry out a thorough analysis of the results of our experiment. To conclude this chapter, we will describe the statistical methodology which was used to answer our research questions.

4.6. Statistical methodology

Choosing an adequate model

Whenever treating a heavy load of data, choosing the optimal statistical model to assess experimental results is rarely a straightforward process, as some parameters such as demographics (like age, sex, university background) and individual preferences can heavily bias the reported results of a poorly fitted test (Brown, 2009). In language learning research, researchers will often use average data of the samples to neutralise individual differences as much as possible (Linck & Cunnings, 2015). While this type of statistical design has its own merits, such as being overall reliable and clear, we wished to consider the effects of individual differences to increase our understanding of latent phenomena in our study. As a result, we were inclined towards using linear mixed models as our statistical asset, in particular for our three initial Research questions.

Indeed, linear mixed models (LMM) are regression models combining the assets of linear models and mixed models, allowing the observer to handle a wide range of variables and scenarios in which observations could be grouped together according to certain criteria instead of being totally independent (Bolker, 2015; Stroup, 2013). Unlike other models used in pre-post-test design research such as ANCOVA, LMM notably allow random effects (variance factors disrupting the assumption of independent

observations) to be taken into account along with fixed effects. While models based only on fixed effects allow us to infer assumptions based on the experiment-specific population, including random effects into a statistical model can often increase its reliability by providing data on population variance without requiring prior averaging (Linck & Cunnings, 2015). The impact of random effects can be especially high when there's a lot of scale (such as a lot of variation in population background, etc.), relatively little data, or an uneven distribution across scales (Bolker, 2015). In such cases, it is considered wise to treat predictor variables as random effects.

Computing random effects

For our statistical models, we treated two variables extracted from Survey 1, namely *Sex* and *University* as random effects due to uncovered factor levels [e.g., we have no data about all *Universities* which could be covered for our study but still wish to make assumptions based on our model], strong variance [e.g., we did observe strong variance in data relative to specific *Universities*], uneven distribution [e.g., we have much more data with some universities than others, but also much more data for *Females* than *Males/Others*] and the fact that these variables are not controlled by our experimental design [e.g. we have no preliminary hypothesis about the effects of *Sex* on score performance]. The lack of distance we have on the effects of these variables implies that they could not be used as fixed effects. Conversely, treating these variables as random effects allows us to infer, to some degree, the potential effects of our experimental design on a larger population not included in our data. For the computation of the effects themselves, we chose to use **random intercept with fixed slopes**, so as to obtain the results of the variation on the outcome variable depending on

random effects – each one of them having its own intercept. Moreover, we have no prior assumption that our random effect variables could have different effects among our treatment groups (CBI/TBLT/NT), which is why the random slopes model should not be used (Midway, 2022).

Computing our fixed effects

As we used a pre-post-test design with independent groups, we also needed to carefully examine the effects of the pre-test when making assumptions. Indeed, pre-test sensitisation should always be considered in pre-post-test designs when possible, especially when we want to control for differentiated tests (Kim & Willson, 2010). Using learning gains (pre-post-test difference) as outcome variable has been shown as risky, as it may reflect the confounding effects of pre-test or post-test rather than intervention effect (Kim & Willson, 2010). Meanwhile, the use of pre-test scores as a co-variate has multiple advantages, such as unbiasing the distribution (as learners with a high pre-test score have lower margin to improve, and vice-versa) and factoring the importance of pre-test in the model's calculations without favouring outliers.

Consequently, we chose to use a residualised change score model to compute the impact of pre-test performance (Valente & MacKinnon, 2017). This model consists of estimating the predicted post-test scores by using the regression of observed post-test scores (Y) on the pre-test scores (X):

$$Y_{i,adjusted} = \beta_0 + \beta_1 \text{Group}_i + e_i$$

where $Y_{i,adjusted} = Y_i - \hat{Y} (= \beta_0 + \beta_1 X_i)$ (Saito, 2020).

For our linear mixed models, using residualised change score effectively means that the pre-test score will be used as a co-variate among other variables for the post-test score regression, instead of being used as a

differential measure with the post-test score. This allows us to obtain regression scores based on adjusted post-test scores accounting for the effect size of pre-test scores (Kim & Willson, 2010).

As all other fixed effects' computation would depend on the research question they'd have to be used in, their integration into the models will be found directly in the adequate description of each model.

In this chapter, we described the way we constructed our experimental design, including pre-/post-tests and surveys, but also the concrete implementation of our experiments as well as the way we chose to analyse all collected data. The following chapter will show the results of the experiments by attempting to answer each research question we described in Section 4.5.

Chapter 5: Experimental data analysis

This chapter will cover the results of the experiments under three sections: a section focused on exposing descriptive statistics, a section thoroughly explaining our mixed models, and a last section detailing learners' responses on a qualitative perspective.

5.1. Descriptive statistics

The key objective of the present research is to measure the effectiveness of two separate methods – CBI and TBLT - in the acquisition of pronominalisation among Japanese university students. As mentioned in Section 3.2.3., pronominalisation as a concept is mainly understood through four dimensions: morphosyntactic forms and meanings, as well as lexical forms and meanings (Purpura, 2013). In the light of Purpura's taxonomy and to assess these dimensions, we constructed two series of tests (one pre-test, one post-test), only separated by a three-week period during which learners were asked to use their attributed game as much as possible. This way, we should be able to relate potential improvements with many variables, such as time on task, completion rate but also the treatment group itself.

While this chapter will be focused on explaining the results of our research, in this first section of our analyses, we will present descriptive statistics of the collected data to have a holistic view of the experiment. We will start with the pre-test and the first survey, then the game variables which were worth mentioning, and finally we will show the data extracted from the post-test and the second survey.

Pre-test scores by university

	University				
	Seinan	Keioh	Fukuoka	Tokyo	Ryukyu
Sample size	<i>n</i> = 40	<i>n</i> = 14	<i>n</i> = 20	<i>n</i> = 8	<i>n</i> = 12
Variable	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Pre-test score E1 (/10)	6.60 (1.68)	6.21 (1.76)	7.35 (1.31)	7.37 (1.68)	6 (1.41)
Pre-test score E2 (/10)	6.95 (2.12)	6.57 (1.34)	6.75 (1.86)	9.00 (1.07)	7.5 (2.50)
Pre-test score E3 (/15)	7.71 (3.01)	7.03 (1.89)	8.00 (2.85)	12.37 (1.64)	7.25 (2.47)
Pre-test score E4 (/10)	4.58 (2.27)	5.03 (1.65)	3.85 (1.67)	6.94 (0.98)	4.96 (1.60)
Pre-test score E5 (/5)	1.46 (1.37)	0.82 (1.08)	1.77 (1.01)	3.37 (0.92)	1.21 (1.19)
Pre-test total score (/50)	27.30 (7.49)	25.68 (6.10)	27.72 (5.82)	39.06 (4.32)	26.92 (7.53)

Table 30: Pre-test scores by university and by exercise type

Table 30 shows differences in pre-test scores and specific exercises scores depending on the universities. Noticeably, the overall mean of pre-test scores was very close for all universities but one (Tokyo University). Interestingly, E1 results were the most balanced across the board, which could hint that it had a lower degree of difficulty. On the other hand, most learners performed poorly for E5 – which required substantial understanding of pronominal structures, as the only task requiring extended receptive and productive knowledge -, the exception being students from Tokyo University. With such a high score achieved in the pre-test, we will be interested to see any improvement with this specific group of students.

Learners preliminary survey

Learners biodata: Sex

Sex	Male	Female	Others
Sample size	27	67	0

Table 31a: Learners' sex distribution

Learners biodata: Age

Age	18	19	20	21	22	23	61
Sample size	1	36	36	18	1	1	1

Table 31b: Learners' age distribution

Learners biodata: Class

Class	1	2	3	4
Sample size	4	44	39	7

Table 31c: Learners' class distribution

Learners preliminary attitudes and perceptions

	University				
	Seinan	Keioh	Fukuoka	Tokyo	Ryukyu
Sample size	<i>n</i> = 40	<i>n</i> = 14	<i>n</i> = 20	<i>n</i> = 8	<i>n</i> = 12
Variable	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<u>Experience</u> ($\alpha = .53$)	2.41 (0.76)	2.68 (0.91)	2.29 (0.88)	3.04 (0.54)	2.57 (0.68)
<i>Experience with pronominal verbs</i> ($\alpha = .85$)	1.85 (0.95)	1.82 (0.83)	2.19 (0.91)	3.43 (0.73)	2.14 (0.75)
<i>Experience with mobile games</i> ($\alpha = .51$)	3.25 (1.42)	3.97 (2.00)	2.45 (1.47)	2.45 (0.79)	3.21 (1.63)
<u>Importance</u> ($\alpha = .63$)	5.4 (0.82)	5.93 (0.75)	4.97 (0.82)	5.34 (0.61)	5.6 (0.87)
<i>Importance of pronominal verbs</i> ($\alpha = .72$)	5.31 (1.00)	5.82 (0.97)	5 (1.08)	5.27 (0.92)	5.97 (0.98)
<i>Importance of mobile games</i> ($\alpha = .44$)	5.54 (1.04)	6.1 (0.87)	4.93 (0.92)	5.45 (0.49)	5.04 (1.68)
<u>Interest</u> ($\alpha = .44$)	4.68 (0.93)	5.4 (0.83)	4.81 (0.89)	4.82 (0.84)	4.58 (0.90)
<i>Interest in pronominal verbs</i> ($\alpha = .48$)	5.32 (1.49)	6.1 (1.07)	5.71 (1.26)	5.95 (0.83)	5.67 (1.09)
<i>Interest in mobile games</i> ($\alpha = -.12$)* ¹⁶	4.25 (1.19)	4.93 (1.18)	4.21 (1.33)	4.07 (1.16)	3.86 (1.35)

Table 31d: Learners' preliminary attitudes and perceptions by university (Likert-scales of 7). Variables in italics are sub-variables of underlined variables.

¹⁶ Considering that this metric was used to measure whether learners prefer one type of game or the other, this negative Cronbach is actually a significant result – a negative α shows that they will usually have a preference (Sijtsma & Pfadt, 2021).

Table 31a, Table 31b, Table 31c present extracted data regarding learners' Sex, Age and Class. We can see that most learners in the study were Female, with a major Age/Class distribution ranging between 19-20 and Class 2-3. Indeed, French classes in Japanese universities are typically more popular with women.

Table 31d shows a tendency for all learners to feel less confident with the use and/or understanding of pronominal verbs. There was, however, a stronger degree of perceived experience level with the target concept among Tokyo University students. Regarding the level of experience with mobile games, learners from Keioh University were more prone to answer positively. As for the level of perceived importance and interest for either pronominal verbs or mobile games, scores were more balanced across the board, albeit there was significantly less interest expressed for mobile games than for pronominal verbs. A limitation of these measures is that Cronbach's α , despite being much improved compared with the pilot experiment, were not significant for several variables (such as experience/importance with mobile games, and the whole interest macrovariable). Still, α were very strong for items such as experience with pronominal verbs ($\alpha = .85$) and importance of pronominal verbs ($\alpha = .72$). Consequently, the accounting of these metrics in our analyses will have to be carefully weighted.

Vocabulary size test

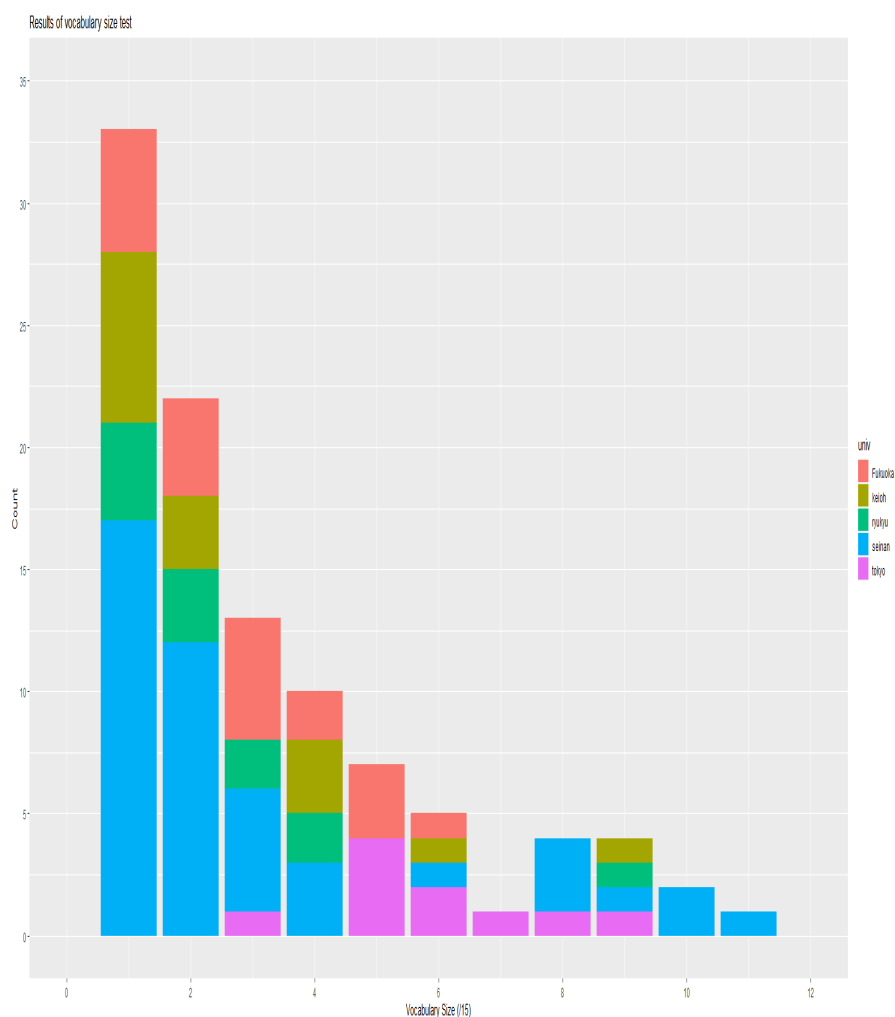


Figure 22: Results of vocabulary size test, sample by university

The results of the vocabulary size test showed that while the overall mean score was relatively low (3.02), there was a lot of variance based on home universities (averages ranging from 2.55 to 5.9). None of the learners scored over 12 out of the maximum 15.

Game data

Time on task (game time)

Time on task (min)	0	>0	≥ 15	≥ 30	≥ 60	≥ 120
Sample size	16	30	7	14	13	14

Table 32: Learners' game time distribution

As game time was a very precise measure (virtually unique to each learner), instead of reporting all the numbers, we decided to show them in 6 time-frame categories. For our games working at 60 frames per second, 3,600 frames equals 1 minute; as such, we converted all our data into minutes for the sake of readability and used these values in our models. It should be noted that these 6 categories were **NOT** used in our models; they are merely used to give an overall view of game time.

Time on task included the sum of game time for all relevant playthroughs (with a minimum game time of 5 minutes). The reason we included all playthroughs was to cover for the potential effects of replayability, which could especially be useful for TBLT learners – as the game encourages new playthroughs due to the ability of choosing game difficulty.

We can see that a good amount of learners did play the game they were attributed to, as only 16 learners (out of 94 learners) did not bother launching the game. Meanwhile, 30 of learners who did play the game did not invest more than 15 minutes in it. 14 learners played for more than 120 minutes. The mean *Time on Task* was of ~53 minutes.

Game completion

Completion (%)	0	>0%	$\geq 20\%$	$\geq 40\%$	$\geq 60\%$	$\geq 80\%$	100%
Sample size	16	26	18	4	3	13	13

Table 33: Learners' game completion distribution

Game completion is a variable we added to check player investment in the games. Indeed, both games are characterised by a lack of forced clock, which means that time on task would increase even when idle. As such, reaching specific milestones in the game would increase a value called Completion, up to 100%. While the increase in completion was, by nature, somewhat arbitrarily chosen, it still has its use to control specific effects (such as players who spent a tremendous amount of time in game without progressing, or players skipping sections to reach a high degree of completion at abnormal game time).

From these descriptive statistics, we can see that the 16 players who did not launch the game are, indeed, marked with 0% completion rate. Interestingly, many more players either reached $\geq 80\%$ or 100% completion rate than $\geq 40\%$ or $\geq 60\%$, which may hint that most learners either stayed at earlier stages of the games or invested heavily into them, with only a few stopping midway through.

Verbalisations and Final Task attempts

Verbalisations were a variable specific to the CBI game, while Final Task attempts were tied to the TBLT game. Unfortunately, the sample number was very low (respectively $n = 4$ and $n = 6$), which means statistical uses of these metrics were too limited. The lack of success of the *Verbalisations* may have to do with the fact that they were locked behind game completion, as an

optional activity; as such, learners who completed the game might not have felt the need to practise further in a bonus-type content. The lack of data for *Final Task attempts* is more concerning, as attempting the final task of the TBLT game was supposed to be the target objective. From what we had heard from learners' feedback, it would seem that many of them did not understand how to access this section; as such, this points towards a limitation in the game's ergonomics, which should be improved for further development.

Post-test scores by university

	University				
	Seinan	Keioh	Fukuoka	Tokyo	Ryukyu
Sample size	<i>n</i> = 40	<i>n</i> = 14	<i>n</i> = 20	<i>n</i> = 8	<i>n</i> = 12
Variable	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Post-test score E1 (/10)	7.00 (2.06)	6.00 (2.35)	6.55 (1.64)	9.00 (0.93)	7.17 (2.33)
Post-test score E2 (/10)	6.63 (1.64)	6.43 (1.28)	5.90 (1.94)	8.75 (0.71)	6.83 (0.94)
Post-test score E3 (/15)	8.00 (2.88)	6.21 (1.93)	7.62 (3.46)	11.44 (2.19)	7.62 (2.25)
Post-test score E4 (/10)	4.75 (2.16)	3.36 (2.72)	4.32 (2.71)	8.50 (0.89)	5.54 (2.47)
Post-test score E5 (/5)	2.44 (1.41)	0.50 (1.18)	1.75 (1.43)	3.65 (0.79)	2.87 (1.45)
Post-test total score (/50)	28.81 (7.53)	22.50 (6.23)	26.15 (7.82)	41.31 (3.52)	30.04 (6.28)

Table 34: Post-test scores by university and by exercise type

Table 34 shows differences in post-test scores and specific exercises scores depending on the universities. While on average, Tokyo University students remained as the top performers for the post-test, considerable differences in means appeared for other universities, with Keioh University and Fukuoka

University students performing worse than for the pre-test (minus 3.18 and minus 1.58 points on average). Meanwhile, we observed improvements in test scores for students from Seinan (plus 1.51 points on average), Tokyo (plus 2.25) and especially Ryukyu (plus 3.13).

Learners feedback survey

	University				
	Seinan	Keioh	Fukuoka	Tokyo	Ryukyu
Sample size	$n = 40$	$n = 14$	$n = 20$	$n = 8$	$n = 12$
Variable	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<u>Ease of use</u> ($\alpha = .91$)	3.80 (1.43)	4.72 (1.49)	3.96 (1.19)	4.10 (0.73)	4.59 (1.27)
<u>Usefulness</u> ($\alpha = .95$)	4.45 (1.37)	4.26 (1.69)	4.46 (1.17)	4.56 (0.66)	4.70 (1.64)

Table 35: Learners' perceived ease of use and usefulness by university (Likert-scales of 7)

The second survey, carried out the same day as the post-test, was mainly used to quantify 2 variables with 6 items each: perceived ease of use and perceived usefulness. With the Cronbach's α being so high, the items were idealistically close and representative of how easy the games were to get into and to navigate in (*Perceived ease of use*), and of how useful the experience with the games was (*Perceived usefulness*). For this survey, we can see that the Likert scores did not deviate much from each other depending on universities.

Pre-post-test scores by treatment

<i>Variable</i>	CBI (n = 38)		TBLT (n = 40)		NT (n = 16)	
	Mean (SD)	Cohen's <i>d</i>	Mean (SD)	Cohen's <i>d</i>	Mean (SD)	Cohen's <i>d</i>
Pre-test	28.07 (7.14)	0.19	28.68 (7.69)	0.04	26.75 (7.85)	0.17
Post-test	29.46 (7.65)		28.95 (7.92)		25.25 (10.11)	

Table 36: Pre- and post-test means evolution by treatment

Table 36 shows the evolution of scores by condition, while also mentioning the effect size via Cohen's *d*. After the experiments ended, the average score mean decreased for the NT condition, slightly increased for the TBLT condition and increased even more for the CBI condition. Due to the decrease in mean score for the NT, we can see that Cohen's *d* metric actually fails to cover the whole effect of the treatments, as it doesn't weight for task complexity (Kraft, 2020). As we used different items between the pre- and the post-test, it is important to note that either test might have been perceived as more difficult than the other, which decreases the reliability of effect size measurements. Furthermore, Cohen's *d* does not cover the effect of **game metrics** inside each condition and considers every individual to receive the same treatment, which is why we will use more elaborate models in order to check on the effects of our experiments. Indeed, as we'd seen previously, there was a lot of variance in time on task and other game metrics.

While throughout this section, we shed light on tendencies relative to universities and other aspects, we chose to avoid interpreting this data in a nutshell, and rather focus on how these variables interacted with each other. As such, the following section will present the results of our linear mixed models, so as to attempt to answer each research question we focused on.

5.2. Linear mixed models

This section will present statistical analyses of learning gains respectively for CBI and TBLT, as the main effect which we are interested in. In particular, we will show the results of the Linear Mixed Models we have used for residualised post-test scores depending on **the investment in the games**, on the whole test (RQ1), then for each type of exercise (RQ2). We will finally deliver an analysis based on other variables which could have impacted on post-test scores, such as **demographics and motivational variables** (RQ3).

Research question 1: Results and effectiveness

♦ Computation

We fitted a linear mixed model to predict the post-test score, as our dependent variable. The model included learners' *sex* and *university* as random effects. The fixed effects initially included were the following:

Treatment (CBI, TBLT, NT), considered as our mediator variable for:

Time on Task [Time on task including playthroughs with 18,000 frames (5 minutes)];

Days on task [Number of days the game was played with 18,000 frames (5 minutes)];

Completion rate;

(Use of *verbalisations*);

(Attempts at *the final task*),

In which *Days on task* and *Completion* interacted with each other as well as with *Time on Task*. Notably, and after a few attempts at including them in our models, we decided to leave out *Verbalisations* and *Final Task Attempt* from our formulae, as the low number of samples could lead to biased results.

As is required when using residualised change score models, *Pre-test score* was used as co-variate, alongside with *vocabulary size*, as it is likely that the overall initial linguistic knowledge as well as the initial understanding of the target concept would influence the post-test results (Bibauw et al., 2019).

Table 37 shows the results of the mixed-effects linear regression on the post-test score. The model's explanatory power of variance is very good, as fixed effects alone have a powerful effect on post-test results (marginal $R^2 = .56$; conditional $R^2 = .71$ ¹⁷).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept) ¹⁸	11.75	3.69	[-91.01, 124.89]	3.184	0.003 **
Pre-test Score	0.59	0.08	[0.43, 0.75]	7.148	<0.001 ***
Vocabulary Size	0.89	0.24	[0.42, 1.36]	3.715	<0.001 ***
TBLT: Time on Task	1.40	0.48	[0.79, 6.35]	2.949	0.004 **
CBI: Time on Task: Completion	0.20	0.08	[0.05, 0.35]	2.621	0.01 *
CBI: Time on Task: Days on Task	0.51	0.23	[0.06, 0.95]	2.240	0.03 *
Random Effects			Estimate	SD	Grouping
University			10.13	3.18	5
Sex			0.21	0.45	2
Residual			19.62	4.43	-

Table 37: Linear Mixed Modelling for RQ1 (selected effects)
[SE = standard error; CI = confidence interval]

◆ Interpretation

Significance in linear mixed models is mostly dependent of the hypotheses; as a rule of thumb, researchers choose to take t values above 2.0 as statistically significant (Linck and Cunnings, 2015). The significant intercept in our model implies that when all predictor parameters are at 0, the

¹⁷ Marginal R^2 only considers the weight of fixed effects, while Conditional R^2 includes the ponderation of random effects

¹⁸ (Intercept) = dummy variable constructed with all parameters at 0

expected outcome variable is significantly different from 0. Since factorial variables such as *Treatment* do not have a natural zero, the model considers using a dummy intercept instead to interpret the effects of other variables, when possible. For that matter, we also attempted to use contrast coded fixed effects (turning *Treatment* values into numeric variables) to test for convergence issues (cf. Linck and Cummings, 2015); however, as the results were practically similar, we chose to keep the three-way factor format for convenience, so that coefficients were calculated in reference with the theoretical dummy value as basis. Without much surprise, the largest fixed effect comes from the *Pre-test score*; indeed, learners having a good understanding of the concept of pronominal verbs prior to the experiment would naturally keep performing well in the post-test (Kim and Willson, 2010). As a result, *Post-test score* was significantly influenced by pre-test performance, proving that the tests were designed to be extremely similar. Meanwhile, the effect of *Vocabulary size* was also significant, as it is usually known as a good indicator of learners' overall linguistic abilities (De Wilde et al., 2020; Milton and Treffers-Daller, 2013; Noreillie et al., 2018).

A visualisation of the relationship (Figure 23) between the *Pre-test* score and the *Pre-post score change* shows that learners with better *Pre-test* score did not necessarily improve more than others; on average, they would stay around the same score from about $\geq 30/50$ in the pre-test (averages showed learning gains of -0.29 for CBI learners, -0.50 for TBLT learners but -3.88 for NT learners). This could hint that learners who already had a relatively high score in the pre-test did not improve much, either because they did not learn much, because they did not play their assigned game anyway or because they considered the post-test as more difficult.

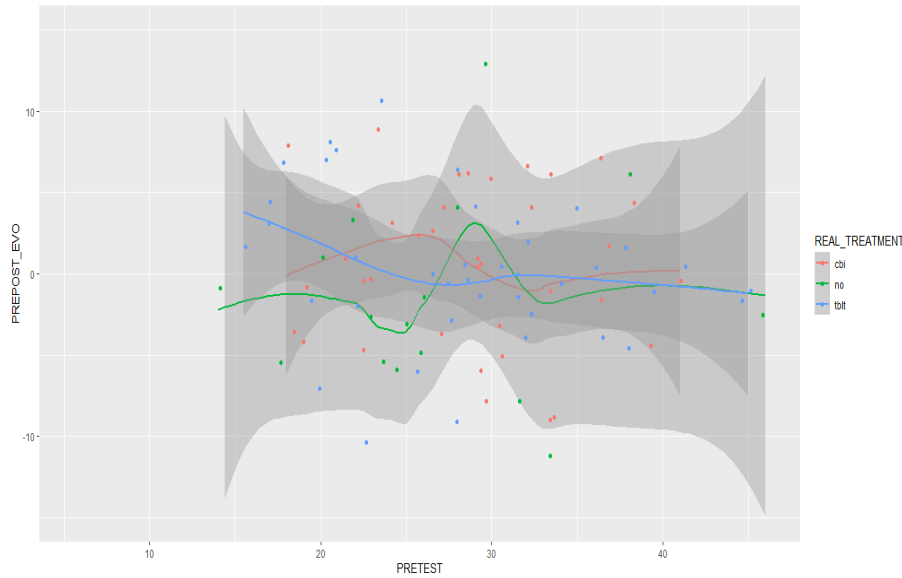


Figure 23: Relationship between Pre-test score and Pre-post score change by treatment

Meanwhile, there was more variance in learning gains for learners with *Pre-test* score $<30/50$. These differences were highly dependent on groups, as for instance, CBI learners with a *Pre-test* score $<30/50$ had a mean learning gain score of +2.22, while it was only +0.85 for the TBLT group and -0.71 for the NT group. However, it should be noted that these measurements, while informative about the overall impact of pre-test scores on learning gains, **do not account for game investment**, which is essential when considered the real benefits of the experiment.

Effects of game investment on learning gains

Regarding fixed effects specific to conditions which included game investment, which is to say measurements relative to the experimental treatment itself, a few observations were made with the model. While game *Completion* rate did not have enough statistical power to account for

significant variance by itself, its interaction with Time on Task was significantly positive for CBI learners. This correlation implies that **time spent playing the CBI game was useful in reaching higher post-test scores, but only when game completion was growing in parallel.**

The correlation between *Time on Task* with *Days on Task* for the CBI condition also helps in understanding that the time spent playing the CBI game was most useful when paired with spread sessions, rather than focused playtime. This effect of distributed practice is in fact not that surprising; in an experiment led with 2 groups of learners with respective instruction time of 3 days and 14 days on the conceptualization of English past tenses showed that retention time was much better for the more spread educational setting (Bird, 2011). While other studies showed similar positive effects of spread treatment on retention (Cepeda et al., 2006; Rohrer and Pashler, 2007), Donovan and Radosevich (1999) warn that the effect of distributed practice is mostly dependent on the complexity of the task. As *Time on Task: Days on Task* was significant for CBI in particular, we could say that it effectively benefited from distributed practice, while **time spent on the TBLT game improved scores independently of the spread in play sessions.**

Although CBI learners improved significantly when several game investment metrics converged, TBLT learners received a much more linear of *Time on Task*, as shown in Table 37. As it was described in Section 3.2, even if it contained several narrative assets, the TBLT game valued immediateness and replayability, as players were quickly able to choose a difficult level and explore freely the different parts of the game; meanwhile, the CBI game started more slowly by setting up characters and story, which in turn could bias the linear effects of Time on Task. Hence, time spent on the games was significantly helpful in itself, albeit in different forms. Figure

24 shows how improvements in scores were influenced by game time, highlighting the linear effect of *Time on Task* on TBLT learners especially:

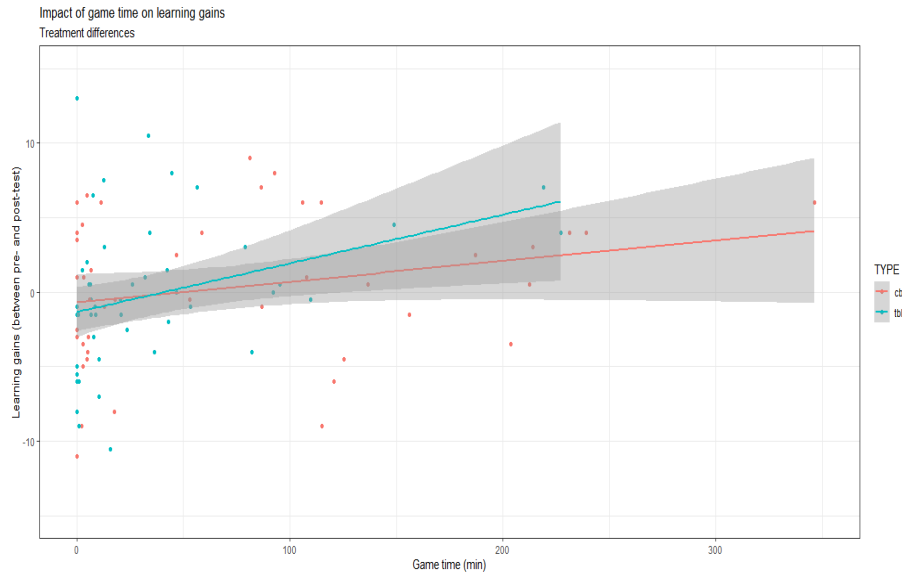


Figure 24: Impact of Time on Task on learning gains (RQ1)

When taking into account *Completion* and *Days on task*, we can see that the investment in both games has similar effects, although the low number of sample for higher *Time on Task* might skew the overall impression:

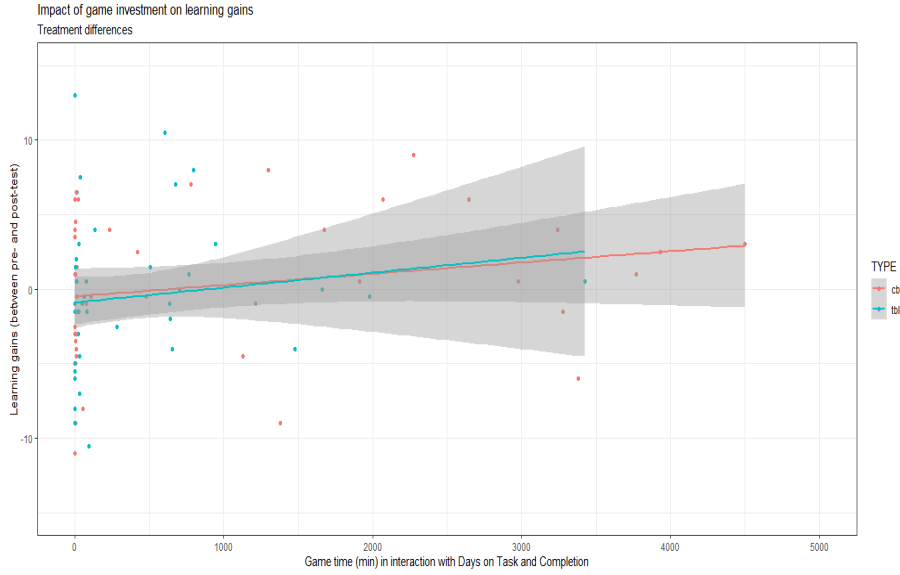


Figure 25: Impact of interacting Game metrics on learning gains (RQ1)

Figure 25 shows that the CBI group had more stable results when taking all game investment parameters into account, while the TBLT group had more widely spread learning gains (as shown by confidence intervals cones).

Finally, we will address the weight of random effects on the model's residual variance. Random effects moderately helped in explaining the residual variance, as calculated by the formula

$$\hat{\sigma}_i^2 = \hat{\sigma}_r^2 / (\hat{\sigma}_e^2 + \hat{\sigma}_r^2)$$

where $\hat{\sigma}_r^2$ is the variance component, $\hat{\sigma}_e^2$ is the residual variance and $\hat{\sigma}_i^2$ is the random effect variance (cf. Bates et al., 2015; Bolker, 2015). *University* was a strong parameter explaining around 34% of the residual variance of the model after fixed effects. Meanwhile, *Sex* barely had any impact on the residual variance. These results show that the differences in scores not weighted by our fixed effects were partially related to learners' home university, but also with other individual factors the model did not cover.

◆ Discussion

According to Nakamura et al. (2021), time on task is a strong indicator of behavioural engagement, especially when it is not forced by an educational context. As learners engage in tasks for longer periods, they are usually expected to improve, making time on task one of the most important predictors of learning (van Gog, 2013). Yet, the relationship between time and learning is not always linear, as there is always a "[...] *need to identify circumstances in which increasing time on task yields tangible benefits for learning*" (Godwin et al., 2021: 515). Indeed, other factors are usually at stake when time on task is correlated with learning gains, such as the quality of the delivery of the instruction or individual factors, such as motivation and interest (Godwin et al., 2021). The linear mixed model we computed for RQ1 demonstrated that investment for both CBI and TBLT was statistically correlated with improvements in scores, either directly (for the TBLT game) or in interaction with other metrics (for the CBI game). The fact that the games were successful at enhancing learners' ability to use and understand pronominal verbs was a prominent finding of this analysis, especially because it might also have hinted learners' overall interest in the games - as they were never **required** to play in the first place.

Research question 2: Results and effectiveness

For the second research question, we fitted our mixed models in the same fashion as for the overall post-test score, but we changed the dependent variable to correspond to the post-test score of **each specific exercise**. Meanwhile, the co-variate *Pre-test* was replaced by the matching exercise score from the pre-test. For these specific models, we specifically chose to highlight the effects of *Time on Task* for the TBLT condition, as well as *Days on Task* and *Completion* interaction for the CBI condition, to verify whether

these game investment metrics were correlated in the same way for specific exercises as they were for the total *Post-test* score.

Post-test Exercise 1 (E1) score analysis

♦ Computation

The first exercise was a Judgment/Recognition task based on selected responses. Higher scores would usually correlate with higher receptive knowledge (Purpura, 2013).

The model's explanatory power was fairly modest (marginal $R^2 = .38$).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	19.21	19.75	[-19.50, 57.92]	0.973	0.33
Pre-test E1 score	0.42	0.13	[0.17, 0.67]	3.322	0.001 **
Vocabulary Size	0.09	0.09	[-0.08, 0.26]	1.042	0.30
TBLT: Time on Task	0.64	0.52	[-0.38, 1.66]	1.225	0.22
TBLT: Time on Task: Completion: Days on Task	0.02	0.01	[-0.003, 0.05]	1.712	0.09
CBI: Time on Task: Completion	0.05	0.03	[-0.01, 0.10]	1.740	0.08 .
CBI: Time on Task: Days on Task	0.13	0.08	[-0.03, 0.29]	1.589	0.12
Random Effects			Estimate	SD	Grouping
University			1.098	1.048	5
Sex			0	0	2
Residual			2.645	1.626	-

Table 38: Linear Mixed Effect Modelling (RQ2, E1)

◆ Interpretation

Table 38 shows notable effects on the Post-test Exercise 1 residualised change score. As expected, Pre-test Exercise 1 (*Pre-test E1*) score had the most impact. *Vocabulary size* was not a significant parameter in this model. As this first exercise tested receptive knowledge in the form of selected responses, the overall lexical knowledge might have had less impact than for more complex tasks, as they "*fail to elicit responses capturing dynamic and complex understandings of the resources needed for communication*", albeit "*useful in situations where the goal is to observe the examinees' receptive knowledge of isolated language features [such as pronominalisation]*" (Purpura, 2013: 16).

The interaction between game investment variables and the Post-test E1 score showed no significant relationships. A couple interactions were indeed close to significance, but could not achieve it: *TBLT* with *Time on task: Completion: Days on Task*, as well as *CBI* with *Time on task: Completion*. As a matter of fact, E1 was considered the easiest exercise, as reflected by mean scores being the closest to the top score, averaging 7.09 for CBI and 6.81 for TBLT. Less demanding tasks might also imply weaker effects from treatment, as many learners are able to answer most items correctly in the first place. As such, no significant relationship was pointed out for this exercise. The effects of *Time on Task* (or lack thereof) on learning gains can be visualised in Figure 26:

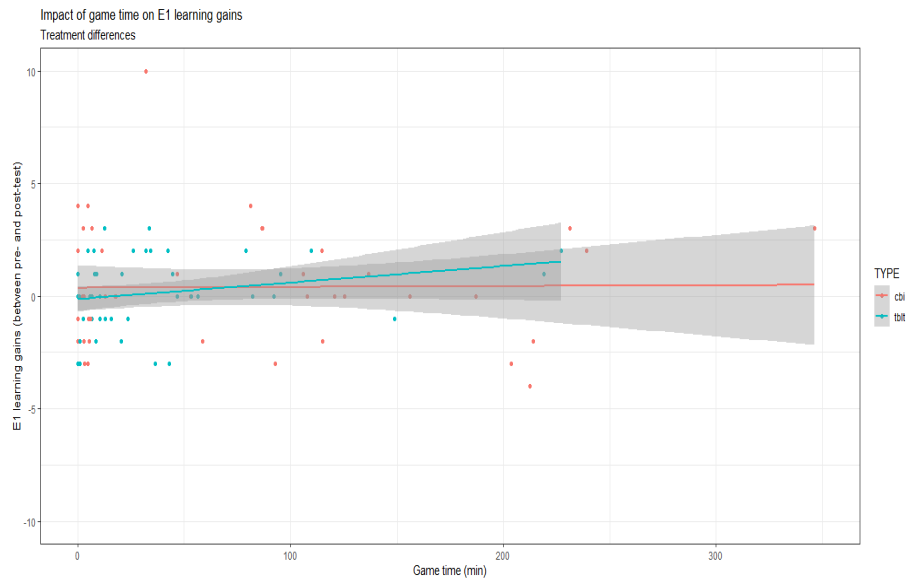


Figure 26: Impact of Time on Task on E1 learning gains (RQ2)

Random effects also had moderate impact on the residual variance, with only *University* explaining about 29% of it, and *Sex* having no impact on the model whatsoever.

Post-test Exercise 2 (E2) score analysis

◆ Computation

The second exercise was an Error-detection task based on selected responses. Just like E1, higher scores should demonstrate higher receptive knowledge (Purpura, 2013).

The model's explanatory power was fairly poor (marginal $R^2 = .32$)

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	-1.03	16.05	[-32.50, 30.43]	-0.064	0.95
Pre-test E2 score	0.25	0.08	[0.09, 0.41]	3.013	0.003 **
Vocabulary Size	0.14	0.07	[0.01, 0.28]	2.079	0.04 *
TBLT: Time on Task	0.90	0.44	[0.05, 1.76]	2.066	0.04 *
TBLT: Completion: Days on Task	0.47	0.29	[-0.11, 1.04]	1.589	0.12
CBI: Completion	1.55	0.86	[-0.14, 3.24]	1.796	0.07 .
Random Effects			Estimate	SD	Grouping
University			0	0	5
Sex			0	0	2
Residual			1.893	1.376	-

Table 39: Linear Mixed Effect Modelling (RQ2, E2)

◆ Interpretation

For the *Post-test E2 scores*, the effect of *Pre-test E2 score* was, unsurprisingly, significant. *Vocabulary size* also had a significant effect on the dependent variable, unlike for E1. Meanwhile Table 39 shows that *Time on Task* had a positive impact on E2 for TBLT users, while it was not the case for the CBI condition. Figure 27 shows that test scores improved specifically for the TBLT group as *Time on Task* increased:

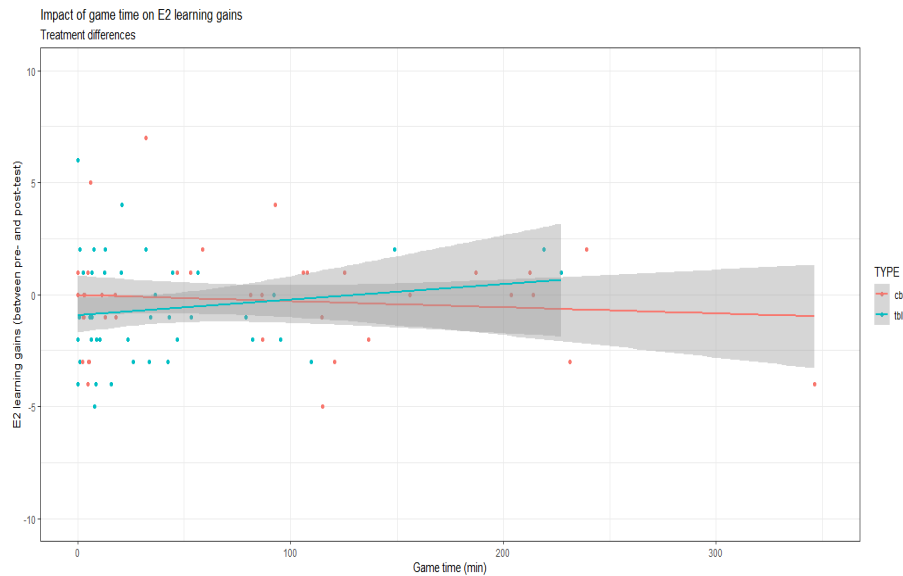


Figure 27: *Impact of Time on Task on E2 learning gains (RQ2)*

Despite the visually noticeable difference with *Time on Task*, the comparison of means between pre-/post-tests E2 scores would suggest that the CBI condition was actually more effective; indeed, average learning gains were negative across the board for E2, going down by -0.55 for NT, -0.75 for TBLT but by only -0.16 for CBI. Yet, for CBI, no interactions were significant on the dependent variable, with the closest being the effect of *Completion* on the dependent variable. As a result, the overall less detrimental effects of CBI on learning gains for E2 were **most likely related to other factors**, such as individual factors.

Surprisingly, the random effects in this model had no weight at all at explaining residual variance; as such, many parameters explaining leftover variance are left unexplained.

Post-test Exercise 3 (E3) score analysis

♦ Computation

The third exercise was a Labelling task with open answers, also known as a closed-ended production task (Mackey & Goo, 2007). This exercise requires productive knowledge just as much as it includes the same type of receptive knowledge of E1 & E2, because learners are asked to ponder the use of the adequate form (pronominal/non-pronominal) and produce it with precision.

The model's explanatory power was strong (marginal $R^2 = .55$).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	5.09	23.32	[-40.62, 50.79]	0.218	0.83
Pre-test E3 score	0.51	0.09	[0.34, 0.68]	5.987	<0.001 ***
Vocabulary Size	0.36	0.11	[0.15, 0.57]	3.402	<0.001 ***
TBLT: Time on Task	0.69	0.63	[-0.55, 1.93]	1.085	0.28
TBLT: Time on Task: Completion: Days on Task	0.03	0.02	[-0.002, 0.06]	1.778	0.08 .
CBI: Time on Task: Completion	0.06	0.03	[-0.01, 0.12]	1.661	0.10
CBI: Time on Task: Days on Task	0.18	0.10	[-0.01, 0.38]	1.865	0.07 .
Random Effects			Estimate	SD	Grouping
University			0	0	5
Sex			0	0	2
Residual			4.03	2.007	-

Table 40: Linear Mixed Effect Modelling (RQ2, E3)

◆ Interpretation

For E3, post-test scores were only significantly impacted by *E3 pre-test scores* and *Vocabulary size*. *Time on Task* was not relevant on E3 scores for either condition. As the first task requiring production skills while still testing receptive knowledge of the target concept, E3 was certainly designed to be more difficult; learners' scores were also the furthest from the perfect score for this exercise (averaging 7.47/15 for NT, 8.03/15 for TBLT and 7.95/15 for CBI). Also, we could see that random effects had no effect on residual variance, which implies that most of the variance was explained by game investment and/or pre-test/vocabulary size. In terms of learning gains, CBI users were also the only group to improve on average, albeit marginally (+0.13, while NT had -0.47 and TBLT -0.27). The interactions shown in Table 40 show that game investment metrics were most likely positively correlated, but lacked statistical power for us to make assumptions. In that regard, Figure 28 allows us to visualise the overall weak effect of *Time on Task : Completion : Days on Task* on the dependent variable:

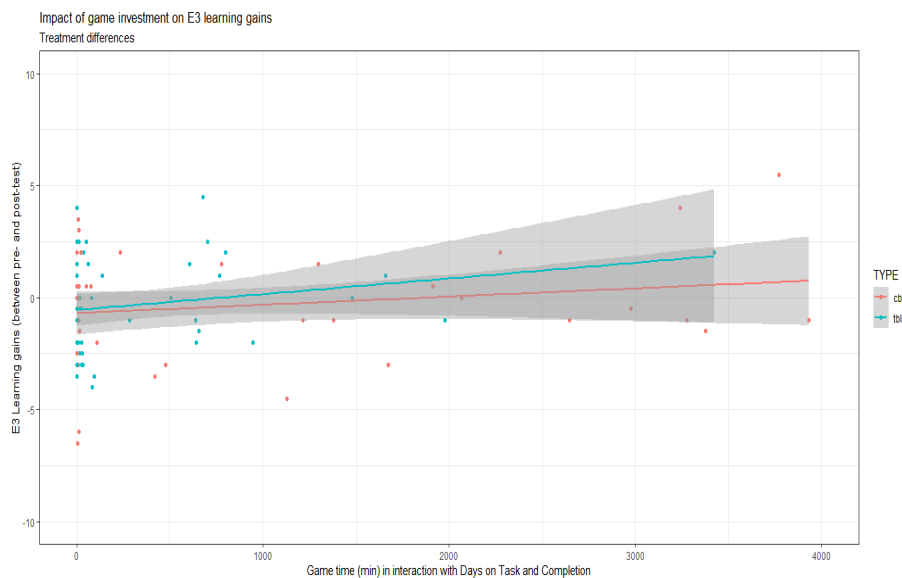


Figure 28: Impact of interacting Game metrics on E3 learning gains (RQ2)

Post-test Exercise 4 (E4) score analysis

♦ Computation

The fourth exercise was a Gap-filler task with open answers. Compared with E3, it adds yet another layer of difficulty as learners are also required to choose the correct verb in a list before using them in the correct form.

The model's explanatory power was strong (marginal $R^2 = .53$).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	8.48	20.78	[-3.22, 49.20]	0.408	0.68
Pre-test E4 score	0.30	0.11	[0.07, 0.52]	2.583	0.01 *
Vocabulary Size	0.39	0.10	[0.20, 0.58]	4.079	<0.001 ***
TBLT: Time on Task	0.69	0.54	[-0.37, 1.75]	1.280	0.20
TBLT: Time on Task: Completion: Days on Task	0.03	0.01	[0.0004, 0.05]	1.989	0.05 *
CBI: Time on Task: Completion	0.07	0.03	[0.01, 0.13]	2.396	0.02 *
CBI: Time on Task: Days on Task	0.18	0.09	[0.01, 0.35]	2.060	0.04 *
Random Effects			Estimate	SD	Grouping
University			1.055	1.027	5
Sex			0	0	2
Residual			2.829	1.682	-

Table 41: Linear Mixed Effect Modelling (RQ2, E4)

♦ Interpretation

For E4 scores, *Time on Task* by itself did not have a significant effect on *Post-test E4* scores. Yet, the evolution of pre-/post-test E4 mean scores suggests that the games had a positive impact on this task too (averaging

-1.03 for NT, +0.38 for TBLT, +0.46 for CBI). Among the fixed effects the model highlighted, the interaction of *Time on Task* with *Days on Task* and *Completion* was significant for the TBLT condition, which implies that **efficient distributed practice** had positive effects on *Post-test E4* scores. Meanwhile, the CBI condition benefited from *Time on Task: Completion* as well as *Time on Task: Days on Task*, emphasising the fact that the CBI game also had positive impact on the dependent variable. However, these effects appeared as very marginal improvements, as shown in the visualization of the most significant interaction (*Time on Task* with *Completion*):

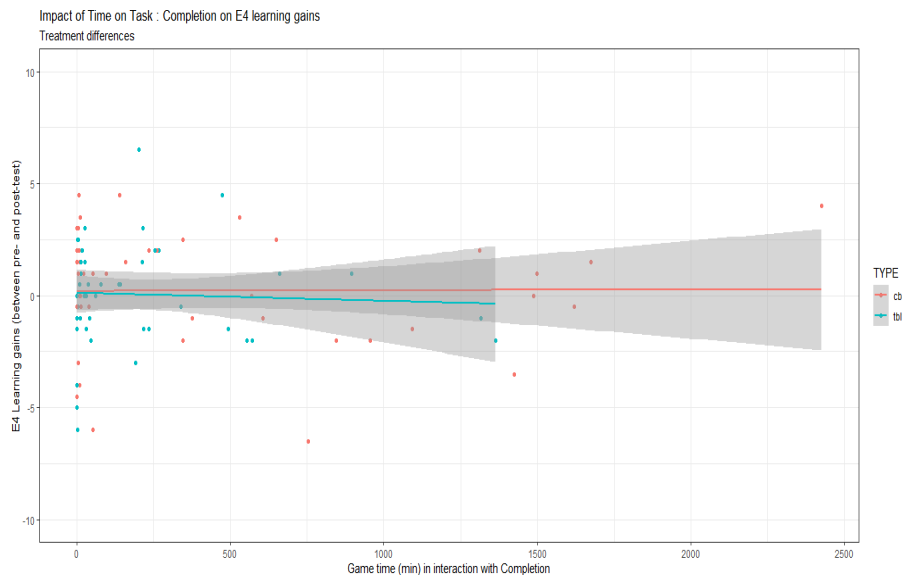


Figure 29: Impact of *Time on Task: Completion* on *E4 learning gains* (RQ2)

Figure 29 shows that visual representations can sometimes appear as misleading due to the lack of accountable data; as such, it is important to always consider the important parameters relative to significance in the model.

As for the random effects, *University* accounted for 27% of the residual variance, which implies that this factor was more decisive in explaining part

of the results compared with E2 and E3. This result appears as intuitive as it is likely that the more complex exercises would get, the more important these differential factors would become, as individuals with more self-esteem or with a higher educational background usually perform better as tasks become more difficult (Judge et al., 2007).

Post-test Exercise 5 (E5) score analysis

♦ Computation

The fifth and last exercise was a Verbalisation-type task requiring learners to justify, with their own words, the use of a pronominal form. Unlike previous exercises, E5 also tests extended production knowledge, in the sense that it allows us to make inferences about learners' full understanding of the targeted grammatical concept (Purpura, 2013).

The model's explanatory power was very good (marginal $R^2 = .64$).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	9.40	10.73	[-1.16, 30.44]	0.876	0.38
Pre-test E5 score	0.47	0.09	[0.29, 0.65]	5.159	<0.001 ***
Vocabulary Size	0.13	0.05	[0.04, 0.22]	2.909	0.004 **
TBLT: Time on Task	0.37	0.28	[-0.18, 0.92]	1.325	0.19
TBLT: Time on Task: Completion: Days on Task	0.02	0.01	[0.004, 0.03]	2.577	0.01 *
CBI: Time on Task: Completion	0.05	0.02	[0.02, 0.08]	3.357	0.001 **
CBI: Time on Task: Days on Task	0.11	0.04	[0.02, 0.19]	2.412	0.02 *
Random Effects			Estimate	SD	Grouping
University			0.2057	0.4535	5
Sex			0	0	2
Residual			0.7645	0.8744	-

Table 42: Linear Mixed Effect Modelling (RQ2, E5)

◆ Interpretation

After the effects of *Pre-test E5 scores*, the impact of *Time on Task: Completion* was the next most statistically significant on the dependent variable. Not only was the interaction significant, it was also relatively strong on *Post-test E5 scores*:

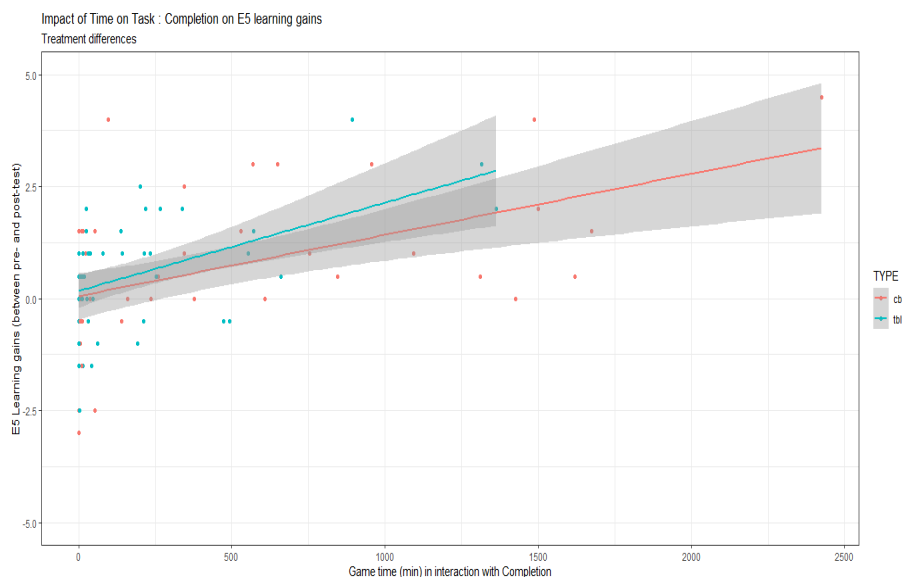


Figure 30: Impact of Time on Task: Completion on E5 learning gains (RQ2)

As an exercise requiring an especially keen understanding of pronominalization, the verbalisation task appeared to be the most connected to game investment metrics. Indeed, a one-way ANOVA revealed that learning gains were significantly different between the treatments for this exercise in particular ($F(2, 91) = [4.061]$, $p = 0.02$). The mixed model shows that it was notably the case when game metrics converged, as *Time on Task* alone was not statistically significant for either condition. Also, it should be noted that the model's residual variance after fixed effects was relatively low. This could imply that the fixed effects we computed were insightful enough in explaining the variance in *Post-test E5 scores*.

◆ General discussion

The analysis of specific exercise scores related to treatment and game investment allows us to confirm our hypothesis for the research question 2. Indeed, we speculated that each game would yield its own impact on

different parts of the test, because each type of exercise focused on specific aspects of the process of conceptualisation. We could see that time spent on the TBLT game was especially beneficial for receptive knowledge, as shown by the significance of *Time on Task* on this condition for E2. The TBLT condition also benefited from the interaction of *Time on Task* with *Completion: Days on Task* for E5, which also tested productive knowledge. Meanwhile, time spent completing the CBI game as well as distributed practice strongly helped for two types of exercises which mostly covered both receptive and productive knowledge (E4 and E5). Interestingly, improvements in *Post-test E1/E3 scores* were not as significantly marked by either group's investment into the games. While no interaction reached a satisfactory significance level for these exercises, it should be noted that they still received positive effects for both TBLT and CBI conditions as demonstrated by the difference in learning gains.

One key finding is that the **models were comparatively better at predicting exercises including production** (E3, E4 & E5), as the higher R^2 demonstrated. This not only confirms that these exercises were, indeed, more difficult - as the results in such tasks are more likely to be affected by game investment and pre-intervention abilities than by unmeasured factors and by chance -, but also puts into perspective the fact that receptive production was better addressed by the TBLT game. As such, the assets of the CBI game on linguistic development and awareness might have been more visible with even more demanding exercises.

Broadly speaking, statistical analyses reported in this section showed, at various degrees, the influence of *Pre-test scores*, *Vocabulary Size* as well as game variables on *Post-test scores*. While the regression models for RQ1 and

RQ2 have shown as though the TBLT game had overall the most potent interactions across the board, learners who were attributed this game did not improve as much on average as for the CBI condition (as shown by mean learning gains in Section 5.1.). As it stands, especially with the little data we have collected, determining what caused this gap in performance can be difficult, as the reasons could be as varied as less attractive game assets, statistical modelling favouring the more direct TBLT approach, lack of personal involvement, etc. All in all, our models allowed us to confirm our two main hypotheses, which is to say that the CBI game offers a slightly stronger effect on learners' performance at the cost of requiring more investment into the game, and yet that both games help in different parts of the process of conceptualisation. These models also offered some perspective on several phenomena we did not have prior hypotheses upon, such as the significant impact of the interaction between *Days on Task* with *Completion* on *Post-test* scores, but also the significant help of higher *Vocabulary Size*. To end Section 5.2., we will showcase the results of our analyses for Research Question 3, which focuses on the effects of attitudes, perceptions, and other survey data on post-test scores.

Research Question 3: Results and effectiveness

◆ Computation

RQ3 focused on the impact of learners' attitudes, perceptions towards games and French, as well as background data which could not be included as random effects on *Post-test* scores. To answer this question, we fitted a linear mixed model to predict the post-test score, as our dependent variable. The model included learners' *Sex* and *University* as random effects. The fixed effects initially included were the following:

Treatment (NT, CBI, TBLT), considered as our mediator variable for:

Experience, Experience w/ pronominal verbs, Experience w/ games;

Importance, Importance of pronominal verbs, Importance of games;

Interest, Interest in pronominal verbs, Interest in games;

Ease of use;

Usefulness;

Age, Class.

For this model too, we used *Pre-test score* and *Vocabulary size* as co-variates, because we wish to extract residualised score changes.

Table 43 shows the results of the mixed-effects linear regression on the post-test score. The model's explanatory power of variance is very good (marginal $R^2 = .53$; conditional $R^2 = .70$).

Fixed effects	Estimate	SE	95% CI	t value	p value
(Intercept)	14.45	6.25	[-67.48, 91.12]	2.310	0.02 *
Pre-test score	0.55	0.08	[-6.41, -1.46]	6.761	<0.001 ***
Vocabulary size	0.58	0.23	[1.03, 6.37]	2.558	0.01 *
Experience with educational games	-0.78	0.42	[-5.71, -0.45]	-1.839	0.07 .
CBI: Ease of use	0.62	0.74	[-0.25, 2.97]	0.827	0.41
TBLT: Ease of use	1.97	0.76	[0.56, 3.02]	2.586	0.01 *
Random Effects			Estimate	SD	Grouping
University			10.22	3.20	5
Sex			8.953e-8	0.0003	2
Residual			17.41	4.17	-

Table 43: Linear Mixed Effect Modelling (RQ3)

◆ Interpretation

For this model, the first notable effect is that neither *Age* or *Class* were significant at any level of the factors. This could appear as counter-intuitive, considering that older learners and/or with more years spent learning French should have a stronger linguistic knowledge, and thus be better at performing at the tests; yet, in that regard, the important variation between universities (*University* accounted for 37% of the residual variance) could have been paramount in cancelling the effects of *Age* and *Class*, as for instance, Tokyo University students were all aged 19-20, but performed much better than older students from other universities on average. Furthermore, the target concept of pronominalisation might not have been better understood by older students, due to how difficult it is to acquire – as was thoroughly shown in Chapter 1.

Table 43 mainly highlights the effects of *Perceived ease of use* on *Post-test* scores. It was especially the case for the TBLT condition, as for CBI, no notable effects were reported by the model. This implies that learners who did feel at ease with their assigned game performed better than others. To be sure, a one-way ANOVA revealed that time spent playing the games was significantly correlated with *Perceived ease of use* ($F(1, 82) = [4.385]$, $p = 0.04$). This analysis shows that learners who played more were most likely feeling comfortable with the game they were assigned to. As suggested in several studies in DGLL (Butler et al., 2014; Reitz et al., 2016; Sandberg et al., 2014), games perceived as less accessible might prevent learners from being successful, which can lead to reduced gains; as such, it is not surprising that learners who felt more comfortable with the game engine performed better. Yet, Figure 31 shows, through the visualisation of the regression lines of this interaction, that the positive effects of *Perceived ease of use* on *Post-test* scores was rather mild:

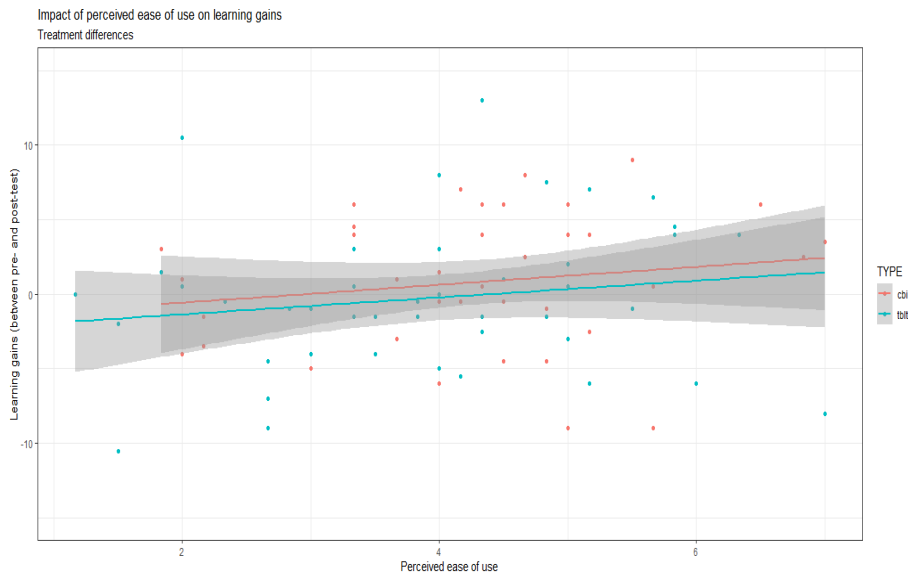


Figure 31: Impact of Perceived ease of use on learning gains (RQ3)

We also noticed that *Perceived usefulness* was not significantly correlated with *Post-test scores*. It is important to note that perceptions (collected with Likert-scale items) are subjective by nature, and as such cannot be considered as precise measurements (Joshi et al., 2015); rather, they are informative about the impression left by the games, as well as the overall interest that learners felt about the experiment.

The mixed-effect modelling analyses carried out in this section attempted to quantify the impact of collected variables on post-test scores for each treatment in this study. Notably, we were able to perceive key differences between the two games, as one was more beneficial for certain tasks than the other, and vice-versa. Our models also showed a significant relationship between game time and the games, albeit in two different forms - *Time on Task* was directly connected with improvements for the TBLT game, while game time was especially significant when interacting with efficient

(*Completion*) and/or spread learning (*Days on Task*) for the CBI game. Finally, we confirmed the importance of *Perceived ease of use* as a contributing factor to post-test scores, in particular for the TBLT game. We can fathom that learners who perceived the games as more accessible were more eager to invest into them, and performed better on average.

In the next section, we will tackle our fourth and final research question: to verify **the impact of each game's typological design on the conceptualisation of pronominalisation**.

5.3. Learners responses analysis

In this section, we will check whether we can detect differences in learners' performance on a qualitative basis by **analyzing their responses to the post-test in comparison with the pre-test**. In particular, we'd like to check whether our typology of pronominal verbs based on diatheses – used in the CBI game - had different effects than the 4-type typology reported by Zribi-Hertz (1987) which was used for the TBLT game. For the sake of reliability, we will only include responses made by users with a **game completion rate of 60% and more**, as with a lower completion rate, more than half of the game's respective typology would not have been studied. As such, these analyses will account for respectively **15 CBI** and **17 TBLT** learners.

The analyses will be delivered exercise by exercise (from E1 to E5), for each test question (pre-test then post-test) and for each covered diathesis. Each sub-section will feature a table mentioning, for every question covering a pronominal verb, the corresponding diathetic category as well as average correctness rates in the identification of pronominalisation, from 0 to 1. Due to the fact that they do not contain any typological information, **questions without pronominal verbs will be left out for this analysis**.

Exercise 1

For the first exercise, 6 correct pronominal forms were used in total for the pre-test, and 5 for the post-test. Only the canonical reflexive and the agentless passive diatheses were used in both versions of E1:

<i>Pre-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	REF	PAS	PAS	-	REC	-	-	-	AUT	REF
CBI Correctness	1.00	.33	.47	-	.47	-	-	-	.80	0
TBLT Correctness	1.00	.18	.53	-	.53	-	-	-	.65	.26

↓

<i>Post-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	-	PAS	REF	-	-	PAS	ANT	-	PAS	-
CBI Correctness	-	.53	.87	-	-	.73	.53	-	.33	-
TBLT Correctness	-	.47	.65	-	-	.59	.53	-	.59	-

Table 44: Evolution in pronominalisation correctness rate, E1
 (REF = canonical reflexive; REC = canonical reciprocal; BEN = canonical benefactive; PAS = potential agentless passive; CON = conversive passive; IMP = impersonal passive; AUT = autocausive; ANT = anticausive)

→ *Canonical reflexive* (Pre-test Q1-10, Post-test Q3):

Canonical reflexives showed an incredibly important gap in correctness rates between the two items from the pre-test, as Q1 tested a well known example of basic pronominal structure, while Q10 tested an abstract use of the verb *se mettre*:

PreQ1 • *Je **m'appelle** François et je suis belge*

PreQ10 • *L'école de ma fille **s'est mise** à distribuer des cahiers gratuitement*

PostQ3 • *On dit que les Espagnols **se couchent** souvent très tard*

As the uses of the canonical reflexives are not easily comparable for these items, we would rather not take for granted that the treatment was the reason

why correctness rates were as high for PostQ3. However, we will keep these structures in mind when analyzing canonical reflexives in other exercises.

→ *Agentless passive* (Pre-test Q2-3, Post-test Q2-6-9):

Agentless passives had different correctness rates depending on the items. In particular, TBLT users had a spectacular evolution (.18 to .59) from Pre-test Q2 to Post-test Q9, which tested the same use of the verb *se trouver*:

PreQ2 • *Elle va [...] à la gare en vélo. La gare **se trouve** près de sa maison.*

PostQ9 • *Le lycée de mon frère **se trouve** près du parc Otori.*

This improvement contrasts with the CBI group, which stayed at the same mean correctness rate (.33). On the other hand, improvement in correctness rate can be seen from Pre-test Q3 (.47) to Post-test Q2-6 (.53, .73) for CBI, which tested structures associated with potentiality:

PreQ3 • *La voiture de mon voisin **s'est vendue** pour 3.000.000 yens [...]*

PostQ2 • *La voiture de mon ami **se conduit** très facilement*

PostQ6 • *Mes nouveaux écouteurs **s'achètent** dans ce magasin*

Meanwhile, correctness rates stayed around the same level for the TBLT groups for these items (from .53 to .47, .59).

Exercise 2

For the second exercise, 6 correct pronominal forms were used in total for the pre-test, and 4 for the post-test. The canonical reciprocal and the agentless passives were the diatheses asked in both versions of the exercise:

<i>Pre-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	AUT	-	PAS	PAS	-	REC	ANT	ANT	-	-
CBI Correctness	.80	-	.27	.33	-	.73	.67	.67	-	-
TBLT Correctness	.82	-	.65	.53	-	.82	.65	.41	-	-

↓

<i>Post-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	-	-	PAS	-	REF	-	-	PAS	REC	-
CBI Correctness	-	-	.80	-	.80	-	-	.93	.93	-
TBLT Correctness	-	-	.76	-	.76	-	-	.76	.76	-

Table 45: Evolution in pronominalisation correctness rate, E2

(REF = canonical reflexive; REC = canonical reciprocal; BEN = canonical benefactive; PAS = potential agentless passive; CON = conversive passive; IMP = impersonal passive; AUT = autocausive; ANT = anticausative)

Immediately, we can notice that the correctness rates were rather high even for the pre-test, for both conditions; perhaps the different question format helped learners identify more easily that the use of pronominalisation was correct for this sentence. Indeed, E2 did not require learners to choose between pronominal/non-pronominal, but to say whether they believe the sentence is correct or not. As a result, learners were more prone to intuitively perceive the correct form compared to E1's format, as the benefits of true-false questions on information retention and lightening of cognitive load compared with more demanding tests is well known (Uner et al., 2022).

→ *Canonical reciprocal* (Pre-test Q6, Post-test Q9):

PreQ6 • *Fiona et Jean se sont rencontrés au lycée*

PostQ9 • *Mes amies se sont rassemblées dans un café*

The structures tested in these sentences were similar, and had an overall high correctness rate for both conditions. While CBI improved between the tests

(.73 to .93), TBLT stayed stable (.82 to .76).

→ *Agentless passive* (Pre-test Q3-4, Post-test Q3-8):

- | | |
|--------|---|
| PreQ3 | • <i>Ce vin se boit avec du poisson</i> |
| PreQ4 | • <i>Au Japon, les voitures se conduisent à gauche</i> |
| PostQ3 | • <i>Une bataille se prépare sur ces terres</i> |
| PostQ8 | • <i>La clé de ma voiture se trouve sur la table</i> |

For the agentless passive diathesis, CBI improved tremendously (.27, .33 to .80, .93), while TBLT had lighter improvements (.65, .53 to .76, .76). PreQ3-4 and PostQ3 were indeed quite similar, as they test the understanding of the potential nature of agentless passives. Interestingly, for PostQ8, we can see that the locative use of *se trouver* was well understood by both conditions – especially CBI learners. The results for this question in particular were inconsistent with E1, as the structure *se trouver* found in E1 was at a stable correctness rate of .33 in both tests for CBI.

Exercise 3

For the third exercise, 8 correct pronominal forms were used in total for each version of the test. The diatheses used in both tests were the canonical reciprocal, the agentless passive, the anticausative and the anticausative diatheses. It is important to keep in mind that the correctness rates mentioned here only reflect the **adequate identification of pronominalisation** in the sentences; while correct spelling was a criterion for our scoring system, it was not useful for this specific analysis, as it does not convey information about learners' understanding of pronominal verbs from the perspective of morphosyntactic forms/meanings (Purpura, 2013). The results were the following:

Pre-test

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Diathesis	PAS	-	-	REC	AUT	-	-	ANT	-	AUT	AUT	-	BEN	-	BEN
CBI Correctness	.80	-	-	.40	.87	-	-	.60	-	.47	.73	-	.73	-	.27
TBLT Correctness	.88	-	-	.53	.65	-	-	.65	-	.53	.65	-	.76	-	.06

*Post-test*

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Diathesis	-	REF	-	PAS	-	REF	-	PAS	-	AUT	ANT	-	REC	-	REF
CBI Correctness	-	.87	-	.93	-	.93	-	.93	-	.87	.67	-	.60	-	.93
TBLT Correctness	-	.94	-	.88	-	.94	-	.82	-	.76	.53	-	.82	-	.76

Table 46: Evolution in pronominalisation correctness rate, E3

(REF = canonical reflexive; REC = canonical reciprocal; BEN = canonical benefactive; PAS = potential agentless passive; CON = conversive passive; IMP = impersonal passive; AUT = autocausive; ANT = anticausive)

→ *Canonical reciprocal/reciprocal* (Pre-test Q4, Post-test Q13):

PreQ4 • *Frank et Anne s'écoutent dans un message vocal*

PostQ13 • *Les partis **se sont consultés** à propos de ce problème*

The canonical reciprocal diathesis in these items was understood better by the TBLT condition, may it be for the pre-test or the post-test. The improvement in correctness rate should be noted for both treatments.

→ *Agentless passive* (Pre-test Q1, Post-test Q4-8):

PreQ1 • *Ce jouet **se range** dans cette boîte*

PostQ4 • *De nouveaux bureaux pour l'entreprise **se construisent** ici*

PostQ8 • *Ces journaux **se vendent** très rapidement*

Regarding the items containing an agentless passive, the rate in the identification of the pronominalisation was already quite high in PreQ1; it stayed around the same height for the two items found in the post-test.

→ *Autocausative* (Pre-test Q5-10-11, Post-test Q10):

PreQ5 • *Je m'adapte facilement aux changements de température*

PreQ10 • *Elle s'arrangera pour venir plus tôt demain*

PreQ11 • *Mon frère s'est illustré au tournoi de tennis*

PostQ10 • *Mon ami s'est égaré dans la forêt quand il était enfant*

Autocausatives were represented by 3 items in the pre-test, with varying correctness rates (from .47 to .87 for CBI, .53 to .63 for TBLT); in the post-test, only one item with an autocausative met great performance from both conditions. In most cases, the lack of direct reflexivity in autocausative sentences did not seem to correlate with increased difficulty for learners.

→ *Anticausative* (Pre-test Q8, Post-test Q11):

PreQ8 • *Son opération s'est passée sans problème*

PostQ11 • *Mon vélo s'est coincé dans un fossé rempli de boue*

Anticausative sentences stayed around the same correctness rate between the two tests, with a small improvement noticed for CBI, while TBLT learners had slightly more difficulty with the post-test structure. Unlike with the agentless passive, anticausatives do not license the potential action of an agent; as such, understanding the nuance added by the pronominalisation can be difficult for learners (Fellbaum & Zribi-Hertz, 1989). Yet, a good amount of learners from both conditions were able to perceive the need to use pronominalisation in these sentences.

Exercise 4

For the fourth exercise, 5 correct pronominal forms were used in total for the pre-test, and 6 forms were used for the post-test. Many errors came from the fact that this task also required learners to identify the adequate verb from a list. Even though only three diatheses were commonly used in both tests,

namely the canonical reflexive, the canonical reciprocal and the anticausative, the improvement between the two tests in terms of ability to identify pronominal verbs was outstanding:

<i>Pre-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	-	REC	REF	-	PAS	-	ANT	-	REC	-
CBI Correctness	-	.53	0	-	.73	-	.27	-	.20	-
TBLT Correctness	-	.24	0	-	.65	-	.12	-	.24	-

↓

<i>Post-test</i>										
Question	1	2	3	4	5	6	7	8	9	10
Diathesis	REC	REF	-	-	AUT	-	AUT	ANT	AUT	-
CBI Correctness	.40	.87	-	-	.73	-	.67	.60	.73	-
TBLT Correctness	.59	.71	-	-	.82	-	.76	.59	.82	-

Table 47: Evolution in pronominalisation correctness rate, E4

(REF = canonical reflexive; REC = canonical reciprocal; BEN = canonical benefactive; PAS = potential agentless passive; CON = conversive passive; IMP = impersonal passive; AUT = anticausative; ANT = anticausative)

→ *Canonical reflexive* (Pre-test Q3, Post-test Q2):

PreQ3 • [...] nous **nous sommes mis** d'accord [...]

PostQ2 • [...] nous avons décidé de **nous promener** en attendant [...]

In the pre-test, the only case of use of the canonical reflexive ended up not being understood at all by learners – the use of the pronominal variant was not perceived by **any** of the 8 learners who had correctly identified the verb (3 for CBI, 5 for TBLT).

In the post-test, the verb was usually correctly identified, but also in the correct form. There seemed to be a big improvement in this category, but we should keep in mind that PreQ3 and PostQ2 compared two fairly different structures, the former being much more abstract than the latter.

→ *Canonical reciprocal* (Pre-test Q2, Post-test Q1):

PreQ2 • *Nous avons donc décidé de **nous revoir** ce matin [...]*

PostQ1 • *Nous **nous sommes retrouvées** à l'aéroport [...]*

The canonical reciprocal saw substantial improvements in the use of pronominal variants for the TBLT condition, as the correctness rate even decreased for CBI learners even though most of them were able to identify the correct verb (Pre-test: 14 for CBI, 14 for TBLT; Post-test: 12 for CBI, 15 for TBLT).

→ *Anticausative* (Pre-test Q7, Post-test Q8):

PreQ7 • *[...] avec des banderoles qui **se déplient** facilement*

PostQ8 • *Un des parfums [...] est tombé et **s'est cassé***

This diathetic category in particular showed great improvement numbers for both conditions (.27 to .60 for CBI; .12 to .59 for TBLT), although part of the difference comes from the better identification of the correct verb (Pre-test: 7 for CBI, 8 for TBLT; Post-test: 13 for CBI, 15 for TBLT).

Exercise 5

Because E5 was a verbalisation task requiring learners to explain the contextual use of pronominal forms, there was no need to identify pronominalisation. As such, the **correctness rates for E5 questions are merely informing about the number of valid definitions** that were given for each treatment. The rates were the following:

<i>Pre-test</i>					
Question	1	2	3	4	5
Diathesis	REF	ANT	PAS	REC	BEN
CBI Correctness	.33	.13	.53	.33	0
TBLT Correctness	.53	.12	.29	.24	.06

↓

<i>Post-test</i>					
Question	1	2	3	4	5
Diathesis	REF	PAS	PAS	ANT	REC
CBI Correctness	.73	.73	.80	.20	.60
TBLT Correctness	.59	.65	.71	.18	.59

Table 48: Evolution in pronominalisation correctness rate, E5
 (REF = canonical reflexive; REC = canonical reciprocal; BEN = canonical benefactive; PAS = potential agentless passive; CON = conversive passive; IMP = impersonal passive; AUT = autocausative; ANT = anticausative)

As we can see, definitions made by CBI were marginally better on average; however, rather than treating these numbers as a rule of thumb on the effects of the games, we believe that focused analyses of learners' open answers is a more adequate method to process the way each condition impacted on the definitions.

For each question, we will show **the evolution in definitions which were considered appropriate (top scores) in the post-test, for CBI then TBLT**; then, we will discuss the potential differences between them.

→ *Canonical reflexive* (Pre-test Q1, Post-test Q1):

PreQ1 • *Je m'appelle Yuki*

PostQ1 • *Je me suis lavé ce matin*

CBI, Learner 1

PreQ1 ▪ (Blank)

↪ PostQ1 ▪ 自分自身を入れるから

[I put myself into the bath]

CBI, Learner 2

PreQ1

- 私が呼ぶという意味になるから。

[Because it means I am the one to call]

↔PostQ1

- 自分自身でお風呂に入っており、
動作主である私はお風呂に入れられる被動者であるため。

[I put myself into the bath so the agent "me" is also a patient put into the bath]

TBLT, Learner 1

PreQ1

- (Blank)

↔PostQ1

- 自分自身を洗うという意味になるから。

[It means that I'm washing myself]

TBLT, Learner 2

PreQ1

- 私は私のことをゆきと呼ぶという意味だから。
私は主語であり目的語でもある。

[Because it means I call myself Yuki. "I" am both the subject and the object]

↔PostQ1

- 自分で自分を洗うため。

[Because I'm washing myself by myself]

As demonstrated by the evolution in correctness rates, learners' responses for the canonical reflexive improved significantly for CBI specifically, although a few TBLT learners also performed better, as shown in the definitions hereabove. Regarding the difference between the two treatments, we could note that one of the CBI learners mentioned the parallel between two semantic roles (agent and patient). The understanding of semantic roles was indeed prominent in the CBI game, as pronominal structures were distinguished by diatheses.

→ *Canonical reciprocal* (Pre-test Q4, Post-test Q5):

PreQ4

- *Paul et Kevin se sont rencontrés au cinéma l'autre jour*

PostQ5

- *François et Marc se sont disputés hier*

CBI, Learner 1

- PreQ4 ■ 二人がお互いに会ったから。会いあったから。
 [Because them met each other. Because they met together]

- PostQ5 ■ お互いがお互いと喧嘩したから。
 [Because each of them is arguing with each other]

CBI, Learner 2

- PreQ4 ■ (Blank)
- PostQ5 ■ お互いに行動を起こしているため。
 [Because they act upon each other]

TBLT, Learner 1

- PreQ4 ■ それぞれがお互いに会うという接触を示しているから。
 [Because it shows a contact consisting of each of them meeting with the other]

- PostQ5 ■ 喧嘩は2人の間で行われたものだから。
 [Because the quarel happens between two persons]

TBLT, Learner 2

- PreQ4 ■ (Blank)
- PostQ5 ■ 喧嘩し合うというニュアンスだから。
 [Because there is a nuance of argueing together]

For the canonical reciprocal, both groups showed tremendous improvements in their definitions, as demonstrated by the similar increase in correctness rates. While different nuances such as “meeting/argueing with each other” or “argueing together” were present depending on learners, it might have simply reflected a choice of words based on typological differences or learners' personality. To be sure, these interpretations were all considered as valid; as such, we could see that the games were very effective on the improvements in the understanding of this diathesis whichever typology was used.

→ *Agentless passive* (Pre-test Q3, Post-test Q2-3):

- PreQ3 • *Le chien de mon voisin s'est fait vacciner contre la rage*
- PostQ2 • *Les voitures japonaises se vendent facilement à l'étranger*
- PostQ3 • *L'école de mon père s'est fait démolir il y a quelques mois*

CBI, Learner 1

- PreQ3 ▪ 犬は自分ではワクチンを打つことはできないから。
[Because a dog cannot vaccinate on itself]
- ↪PostQ3 ▪ 学校は自動的に壊れない。会社などによって取り壊されたから。
[Because a school cannot demolish itself. It was destroyed by a company or something]

CBI, Learner 2

- PreQ3 ▪ 受け身
[Passive]
- ↪PostQ3 ▪ 学校が「壊される」という受け身を取るから。
[Because the school takes the passive "is demolished"]

TBLT, Learner 1

- PreQ3 ▪ ワクチンが犬の体内に入り込み、犬のものとなるため。
[Because the vaccine is put inside the dog's body, it becomes part of the dog]
- ↪PostQ3 ▪ 学校が取り壊されるという受け身だから。
[The school takes the passive "demolished"]

TBLT, Learner 2

- PreQ3 ▪ (Blank)
- ↪PostQ3 ▪ 原因を重視しないものだから。
[Because the emphasis is not on the cause]

The agentless passive was another case of diathesis which was assessed in a much better way in the post-test for both conditions. Not only did the verbalisations gain in sophistication, by using more specific concepts such as 原因(*cause*) or 自動的に(*by itself*), they were also less confusing (cf. TBLT Learner 1). As a result, we did not see a specific effect of one of the two typologies on the understanding of the agentless passive.

→ *Anticausative* (Pre-test Q2, Post-test Q4):

PreQ2 • *Le verre s'est cassé malencontreusement*

PostQ4 • *Le coût de la vie s'accroît de plus en plus*

CBI, Learner 1

PreQ2 ■ グラスは自分から割れるのではなく、
 なにか別の要因があって割れるから。
 [The glass doesn't break by itself, it breaks because on external cause]

↪ PostQ4 ■ 物価は勝手に上がらない。
 私たちのいろいろな要因によってあがるものだから
 [Cost of living does not increase by itself, but due to the effects of our actions]

CBI, Learner 2

PreQ2 ■ 主語が割れたから
 [Because the subject broke]

↪ PostQ4 ■ 理由がわからない
 [Because we don't know the cause]

TBLT, Learner 1

PreQ2 ■ (Blank)

↪ PostQ4 ■ 原因を重視しないものだから。
 [Because the cause is not emphasised]

TBLT, Learner 2

PreQ2 ■ グラス自体が割れるという状態になったから。
 [Because the glass itself entered a broken state]

↪ PostQ4 ■ 何かの影響で物価の変動が起こっているから。
 [Because the change in cost of living is caused by the influence of something]

The anticausative diathesis did not see any significant improvements between the pre-test and the post-test; in fact, most of the learners who did well in the post-test also had a good understanding of the diathesis initially (TBLT, Learner 1 cited hereabove being one of the exceptions). By a matter of fact, the majority of targeted learners were not able to perceive the

importance of the cause - or the lack thereof – in this diathesis, identifying the sentence as a passive one. Surprisingly, the concept of cause was present in two definitions made by TBLT learners, even though it was not mentioned in the game. The relationship between the pronominalisation in these structures and their anticausative nature may be more intuitive than what we had originally thought.

Discussion

Throughout this section, we attempted to portray the differences in the **evolution of the ability to identify pronominalisation depending on each treatment**, so as to verify whether certain diatheses/categories were more difficult to process for one condition than the other. Diathesis by diathesis, the **mean aggregated gains in correctness rates for all exercises**¹⁹ were the following:

<i>Diathesis</i>	<i>CBI mean aggregated gains in correctness rate</i>	<i>TBLT mean aggregated gains in correctness rate</i>
<i>Canonical reflexive</i>	+ .38	+ .14
<i>Canonical reciprocal</i>	+ .14	+ .23
<i>Potential agentless passive</i>	+ .21	+ .19
<i>Autocausative</i>	+ .18	+ .15
<i>Anticausative</i>	+ .16	+ .14

Table 49: Global gains in pronominalisation correctness rate by diathesis

We can see that on average, CBI gains in terms of correctness rates were only slightly better, except for the canonical reflexive, for which they were substantially bigger, and the canonical reciprocal, for which the TBLT

¹⁹ Calculated by averaging the mean correctness rates of all exercise types, for each diathesis, by treatment.

condition improved more. However, the analyses showed that differences appeared according to the concordance of two criteria: diatheses and exercise type. Indeed, in tasks only testing receptive knowledge (E1 and E2), despite performing initially worse on most of questions, CBI learners ended up with a higher correctness rate except for one type of structure (the locative *se trouver* in E1). Meanwhile, E3 and E4 – which also tested limited productive knowledge – saw an overall better recognition rate of pronominal structures for the TBLT condition. Finally, for the verbalisation task testing extensive productive knowledge (E5), the CBI group had the best performances.

These results are somewhat startling, as we had seen in the exercise-based statistical analyses in Section 5.2. that TBLT learners usually benefitted more than CBI learners from game investment for E1 and E2, while the situation was reverted for E3-E5. Yet, the learners targeted for the qualitative analyses hereabove usually invested a lot of time into the games, as their completion rate was over 60%. We could hypothesise that these paradoxical results would originate from differential results in the other questions found in E1 to E4, which did not cover typological information and were thus left out for these analyses. Yet, pre-post test gain scores were only marginally better for the CBI learners with a completion rate over 60% than for their TBLT peers (+2.17 on average for CBI *vs* +2 on average for TBLT), while post-test score means were higher for TBLT in the targeted sample (with completion rate over 60%), at 29.38 instead of 29.30. This shows that **typological differences could not assess the total difference in scores when observed only from the perspective of identifying pronominalisation**; it does not allow us to account for decisive factors such as accuracy in responses, recognition of situations in which a pronominal verb should *not* be used, as well as other sources of mistakes (such as blank responses). Furthermore, as

it was shown in Table 49, the limited number of items treated did not encompass all pronominal structures which could be relevant in both tests – in fact, 3 diatheses were used in one test and not the other, preventing us from making comparisons. Consequently, the results of these analyses should rather be considered as informative than conclusive, especially when compared with the significant effects of game investment metrics on test scores.

Notwithstanding the **lack of important difference between treatments** in correctness rates in regard with the identification of pronominal verbs, the most interesting outcome of the analysis of verbalisations, in particular, lies in the **increased awareness of learners** regarding the functions of pronominalisation **in a pre- and post-test perspective**. According to Jarvis (2011), the acquisition of newly known concepts goes through *conceptual transfer* when grammatical concepts are not fully *congruent* between the first language and a foreign language. This conceptual transfer implies the restructuring of the current conceptual categories to match a new mental state, while most of gains in lexical knowledge do not necessarily cause this restructuring of the mental representations. Pavlenko (1999) cites the example of aphasic persons, who are able to categorise a *mug* as different from a *cup*, even though they would not be able to access these words from a lexical perspective.

In the case of our learners, **the gains in the ability to identify situations in which pronominal verbs should be used were substantial**, compared to prior to the intervention. This result has to be emphasised as the main outcome of the qualitative analyses of this section, especially because it was the case for both experimental conditions – albeit the gains in correctness rates were a bit higher on average for the CBI treatment. This means that

conceptual transfers – as defined by Jarvis (2011) - were made possible with the help of both games, at different levels. As such, it is not an easy task to verify whether one treatment offered better opportunities to conceptualise than the other; as learners progressed through their respective game, most of them were able to increase their understanding of pronominalisation.

5.4. General discussion

In this last section of this chapter, we will summarize how the analyses conducted in Section 5.1. to Section 5.3. helped in answering each one of our Research Questions.

RQ1: Statistical evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, in correlation with game investment metrics

The analyses conducted in Section 5.2. showed that game investment metrics (*Time on Task*, *Days on Task* and *Completion*) were generally correlated with higher test scores. While the impact of *Time on Task* was significant on post-test scores for the TBLT condition, it was also a key factor in the learning gains for the CBI condition when associated with distributed practice (represented by increases in *Days on Task*) and efficiency (represented by increases in *Completion*).

When looking at raw learning gains by treatment, the three conditions did not seem to improve much on average (CBI gained +1.39, TBLT gained +0.28, while NT decreased by -1.5). However, as mentioned before, learners within the same group did not necessarily have the same treatment (as their game experience was individual); tests also included different items, albeit

built upon the same guidelines. As such, the use of other variables, such as comparing learning gains along with *Time on Task* below 30 min/over 30 min, shows how the games generally benefitted both conditions (CBI gains with a *Time on Task* below 30 min: -1.3; over 30 min: +3.15; TBLT gains with a *Time on Task* below 30 min: -1.3; over 30 min: +2.19).

These results mostly demonstrate that learning gains were generally higher for the CBI condition; however, these differences could be mitigated by the fact that CBI learners were more invested into the game than their counterpart (mean *Time on Task* of 120 min vs 76 min for TBLT). In contrast, TBLT learners showed improvements in learning gains for a smaller amount of time invested into the game, while the 16 learners which did not play the games performed worse in the post-test on average.

RQ2: Statistical evaluation of each group's post-test performances for each specific exercise regarding the use and understanding of pronominalisation before and after the experiment, in correlation with game investment metrics

Unlike RQ1 which addressed overall learning gains, RQ2 attempted to see whether the games benefitted more in certain types of tasks than others. Part of the analyses conducted in Section 5.2. were focused on answer this question. To be sure, E1 (exercise 1) and E2 included a series of tasks testing receptive knowledge; E3 and E4 also required of learners to produce forms, and as such also tested limited productive knowledge in contextualised sentences; finally, E5 involved both receptive knowledge and extensive productive knowledge, in the sense that learners were asked to provide full verbalisations of their understanding of a pronominal form.

The results showed that time spent (*Time on Task*) into the TBLT game was statistically correlated with improvements in E1 and E2, while CBI learners' investment in their assigned game was linked with better results in E4 and E5 in particular. These observations allow us to infer that whilst the TBLT game was the most effective on learning gains in tasks focused on receptive knowledge, the CBI game had the best effects for tasks involving both receptive and productive knowledge. Meanwhile, the linear model we used for E3 showed a lack of significance for game investment metrics on test scores.

RQ3: Statistical evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, in correlation with learners survey metrics

RQ3 sought to verify the effects of non-game related variables (attitudes and perceptions towards games and French pronominal verbs, learners' *Class* and *Age*) on learning gains. Notably, we carried out a linear mixed model in Section 5.2. to obtain answers to this question.

The results showed that most non-game related variables lacked significance to be considered as impactful on learning gains. However, *Perceived ease of use* was a stand-out, as learners who perceived their assigned game as easier to use usually had better learning gains (especially for the TBLT condition). In this regard, we were able to retrieve learners' feedback about accessibility problems as part of the data gathered with Survey 2. Some of the feedback received were very insightful regarding the importance of *Perceived ease of use* on their experience with the game and, potentially, on learning gains:

文字が打ちにくかった。

[It was difficult to type inside the game]

出てくるセリフが全部フランス語とは思わず苦戦

[Many lines were in French so it was very difficult]

難しかった。ゲームの使い方は分かったが内容が理解しにくかった。

[It was difficult. I understood how to play but the content was complex]

Designing games to be accessible is paramount in DGLL, because an unfriendly interface or linguistic content perceived as too complex may hinder the players and prevent them from investing further into the game. As a matter of fact, the learners cited hereabove had rather low playtime (below 30 minutes) and performed worse in the post-test. In such cases, the adverse impact of accessibility problems, reflected for instance by a low *Perceived ease of use* value, could be seen immediately and calls for adaptations in further development.

RQ4: Qualitative evaluation of each group's post-test performances regarding the use and understanding of pronominalisation before and after the experiment, from a typological perspective

Finally, RQ4 was focused on the typology-based analysis of learners' responses. In particular, we looked at the evolution in correctness of learners' responses with a *Completion* value over 60% in test items which involved the use of pronominal verbs. The typology we used as benchmark was the diathesis-based typology (cf. Section 2.2.4.2.), so that the responses could be

analysed with more precision. Structures which did not require the use of pronominalisation were excluded, because they could not be informative enough about the impact of the typologies covered in the games (in such cases, misuses of pronominal verbs can hardly be categorised as typological mistakes).

The first and foremost result is that there was an impressive increase in correctness rates across the board. The only category covered by both tests in which TBLT learners had better gains than their counterpart was the *Canonical Reciprocal* diathesis; meanwhile, CBI learners had stronger improvements in correctness rates in the other diatheses – especially for the *Canonical Reflexive* diathesis.

Furthermore, the qualitative analysis of learners' open responses in the verbalisation task (E5) showed that both groups included learners whose quality of definitions improved tremendously between the pre-test and the post-test. Although it was not exclusive to the CBI group, we also noticed a trend for several CBI learners to include the use of grammatical concepts such as semantic roles and syntactic functions in their analyses. This shows important gains in the quality of learners' meta-analysis behind the concept of pronominal verbs, which might be reflected by long-term gains. Unfortunately, we were not able to carry out delayed post-tests, which limits the range of analysis of the results covered by this question in particular.

All in all, this chapter enabled us to answer our questions in a satisfactory fashion. Statistical analyses showed that both games were effective at pushing the comprehension of pronominalisation for different stakes; meanwhile, the analysis of learners' responses hinted that the use of separate typologies in each game did not reveal major gaps in the way learners

handled pronominal verbs, especially as reflected by the improvements in the quality of verbalisations made by both groups. Yet, the fact that the games had a positive impact on the acquisition of grammatical content can be considered a major outcome by itself, as the literature in DGLL seldom reports on experiments tied to grammar learning (Su et al., 2021).

In the next and final chapter, we will portray a summary of the findings of the present work, highlighting its contributions but also its limits; at last, we will attempt to offer some insights on where to improve on the domain for further research, so that we could develop even more efficient ways to assess conceptual knowledge.

Chapter 6: Perspectives and conclusion

6.1. Summary of Findings

All throughout our PhD research rendered visible in the present work, we were motivated to provide insights and methodological answers to a problematic situation regarding the acquisition of pronominal verbs for Japanese university students.

In Chapter 1, we built an overview of these problems in general, but also in the context of Japanese universities before attempting to suggest leads towards potential workarounds.

Chapter 2 was focused on presenting the theoretical tenets behind our main experimental design, which is to say CBI, while reviewing experimental studies up to 2025 revolving around this methodology. We also investigated the target concept of pronominal verbs in order to find a way of explaining it in a systematic way so as to build a SCOPA, notably with the help of the annotation of two corpora, *Sequoia* and *FTB*.

Chapter 3 presented the system design chosen to examine the development of the target concept, namely the use of mobile games. After reviewing the literature regarding the applications of Digital Game-based Language Learning, we detailed how a CBI-based game and a TBLT-inspired game – as our comparison group - were constructed with the help of RPG Maker.

Chapter 4 explained the experimental design used to evaluate students' learning gains. After describing the participants, we clarified the methodological choices which were made to generate online tests and

surveys, so as to verify the development of the target concept in a pre-post experimental design. Finally, we depicted the conduct of two experiments involving the two games we had created and that were carried out in Japanese universities in 2023 and 2024, notably showing the way data was collected.

Finally, Chapter 5 attempted to deliver, through statistical and qualitative analyses, satisfactory answers to the four Research questions that were at the focal point of the present research. The results of our statistical analyses and our diathesis-based analysis helped us understand better how the process of pronominalisation developed with learners in our two experimental conditions. Several aspects have to be taken out of this chapter as the most prominent findings of our research:

a) Conceptualisation goes through different trajectories depending whether learners learn by directly using their first language to assess L2 concepts, or by using their L1 as a mediator to assess L2 content.

Our CBI-based game was a prime example of an activity which attempted to activate L2 concepts directly with learners' L1. Statistical analyses (especially related to RQ1 and RQ2, in Section 5.2.) showed that this type of design could lead to stronger gains overall, especially when assessing tasks requiring productive knowledge along with receptive knowledge, as **all game investment metrics** in the game grew bigger. Meanwhile, diathetic analyses demonstrated that the gains in pure conceptual awareness were comparatively stronger with most diathetic structures, albeit it was especially true for tasks based on receptive knowledge only, as well as for the verbalisation exercises.

On the other hand, our CLT/TBLT-inspired game was representative of a design based on L2 tasks to access L2 concepts, while offering sufficient L1

support to ensure salience (translations/glossary). Statistical analyses showed that **game time investment** in this game especially favoured improvements based on receptive knowledge, while diathetic analyses displayed slightly better gains than CBI in conceptual awareness for tasks related to limited productive knowledge.

As a result, there does not seem to be an outright *better* method among CBI or TBLT to acquire the concept of pronominalisation; it would appear that both trajectories can be beneficial, if only sufficient investment was made by any learner. In fact, our first and foremost hypothesis, which was to say that the CBI game would lead to better results relative to the process of conceptualising pronominalisation, is either debatable or only true by a very small margin, as we have observed that CBI learners did improve more on average than the TBLT group. Yet, the effect of both games on the acquisition of the target concept was undeniable, as learners from both experimental groups strongly outperformed the No Treatment condition on average, as it was shown in Section 5.1.

b) The effects of our diathetic typology of pronominal verbs compared with Zribi-Hertz's 4-categories typology (1987) was only relatively successful.

A major asset of the experimental methodologies tested in the present work was that they covered pronominal verbs through a specific taxonomy. This design choice was made so as to respect the theoretical tenets of CBI and TBLT, as these approaches have their own way of delivering conceptual content. Statistical results could not inform precisely about the gains relative to typological differences, as the games were also different in other aspects; this is why we also conducted a qualitative analysis of the recognition of pronominalisation based on diatheses in Section 5.3. This analysis showed

that for most structures of pronominal verbs - except for the canonical reciprocal diathesis -, CBI learners benefited the most from the typology used in their assessment game. However, the difference in gains was not as big as we would have imagined initially. This implies that both typologies could be used to conceptualise most pronominal structures almost indiscriminately, even though typological effects on a couple specific diatheses could not be verified in this work due to a lack of comparison grounds.

c) Autonomous learning can be useful in SLA contexts, especially when combined with distributed practice.

As a consequence of the success of our games on the acquisition of pronominal verbs, we were reminded of the importance of offering tools for learners to practise autonomously, at their own pace. We were especially pleased by the fact that many learners offered positive feedback on their experience with the games:

解説がわかりやすくて気軽にできるので良いと感じた。

[I found the explanations easy to understand and easy to do.]

説明がわかりやすくて面白かった

[The explanations were clear and interesting.]

代名動詞がどのような時に使われるのかを学習するのにすごく役立ちました！

[It really helped me learn when pronominal verbs are used!]

これまでどのような場面で代名動詞を使うのか曖昧にしていたのですが、ゲーム内で使うときの条件や例文を出して下さったので、すごくわかりやすく勉強することができました。

[I had been vague about in what situations I would use pronominal verbs, but you gave me the conditions and examples of

when to use them in the game, so I was able to study them very clearly.]

キャラクターの設定が細かくて、とても楽しかったです！！

[I really enjoyed the detailed setting of the characters!!]

ゲームをすることによって、簡単に代名動詞の使い方を理解できた。楽しみながら学ぶことができた。

[By playing games, I could easily understand how to use pronominal verbs. I was able to learn while having fun.]

Educational games used in autonomous contexts have the advantage of encouraging learners to commit to decisions while not fearing the consequences of mistakes (Veget, 2018). Furthermore, we wished to guarantee that every student could access the game in their free time, as it allows learners to practise while avoiding the pressure of being confronted by the look of their peers (Ishikawa, 2011). This is why the choice of a mobile game made sense, as in Japan especially, many university students only possess a mobile phone and do not have a personal computer. As the feedback showed, the benefits of the games also resonated with the well-referenced lack of self-confidence of Japanese learners (Kambaru, 2023), improving their ability to correctly identify and interpret subtle characteristics of French pronominal verbs.

Finally, an important asset of autonomous learning through DGLL also comes from its compatibility with distributed practice, as we had seen throughout our statistical analyses that the interaction between *Completion* and *Days on task* was often relevant for better test performances. By nature, mobile games with autosaves will allow learners to spread their learning process over time (*self-regulation*), which in turn is often translated by stronger treatment effects (Cepeda et al., 2006; Rohrer & Pashler, 2007).

d) Learners' biodata do not matter as much in a relatively homogenised context.

In the course of our analyses, we took care to check for the effects of several variables concerning learners' biodata, such as *Sex* and *University* as random effects, and *Age* or *Class* as fixed effects. From the results of our mixed models, we could see that *University* was the only variable which had significant impact on the results' residual variance, with weight depending on the outcome variable. As such, we consider that the games appealed to all learners within our target sample indiscriminately.

e) Other findings: preliminary linguistic knowledge, attitudes and perceptions.

Finally, we should remind that many other variables were significant in the development of learners' acquisition of pronominal verbs. Among the most important variables, the key role of *Pre-test scores* and *Vocabulary size* cannot be undermined. Indeed, as positive as game investment metrics could have been, the fact that preliminary knowledge, understanding of pronominal verbs as well as lexical knowledge were strong predictors of post-test scores. Meanwhile, variables related to attitudes and perceptions were not always relevant (either in a nutshell, or due to poor item selection), although *Perceived ease of use* of the game was a stand-out. As such, researchers in DGLL should always be cautious when designing educational games, so as to proactively make games that allow learners to feel comfortable with in order to garner the best possible results.

f) Overall effectiveness of the experiments relative to the target population.

All in all, these results showed that the experiments carried out were largely

successful, especially when considering that each game's approach on the acquisition of pronominalisation was different but still beneficial. These findings also resonate with the fact that the games were appreciated by the target population sample, which is to say Japanese learners. The analyses carried out in Section 5.3, notably, showed that several learners who had completed a decent portion of their respective games improved significantly in their awareness of pronominalisation. As shown in several studies (Delbarre, 2010; Nakamura, 2004; Soga, 2003) but also in our analysis of French textbooks used in Japan made in Chapter 1, the teaching of pronominal verbs in Japanese universities can be lacking in terms of assessment of meaning, to the point that Japanese learners shy away from using them. As a result, the meaning of pronominalisation itself is rarely addressed, and Japanese students only show strong levels of mastery of this concept at later stages of French acquisition (B2-C1, cf. Section 1.4). Yet, the analyses we carried out demonstrated that Japanese learners who invested in the games were especially able to conceptualise contextualised uses of pronominal verbs in a very satisfactory manner, even though they were most likely at a proficiency level lower than what is usually expected for the acquisition of this concept in particular.

The experiments were also strongly marked by positive feedback from the students. Many learners demonstrated a keen interest in learning pronominal verbs with the games, to the point that the games were still played by some learners after the experiment was concluded (as data continued to reach the game server). While there is also much improvement to be made in terms of design - especially to build the TBLT game as more accessible, our research showed that the use of DGLL in order to teach pronominal verbs was a promising ground. In this regard, rather than proposing only a single game, it

appeared as though combining both games' assets into one could bring the best results, as the games usually helped in acquiring different aspects of the understanding of pronominal verbs.

g) Integration of our work in the body of CBI research.

Finally, it is paramount to reflect on the results of these experiments in the light of the body of CBI research in the domain of second language acquisition. While we believe that we have respected the theoretical tenets of CBI as much as we could, we also think that this work attempted to innovate on several fronts:

- It integrated Digital-based Language Learning in a CBI approach, which had never been done before to the extent of our knowledge;
- This work involved the creation of an alternate condition based on CLT/TBLT with a similar success, yet showing different benefits regarding linguistic development;
- The use of linear mixed models helped in controlling the effects of the module on the acquisition of pronominal verbs in a precise way, and even use non-player residuals as a *post-hoc control group*;
- We also confirmed that verbalisation made throughout actions (already experimented in Serrano-Lopez & Poehner, 2008) is a potent way of constructing conceptual knowledge, as the free definitions in E5 demonstrated a high level of awareness.

We also believe that the important participation rate in a voluntary experiment, the higher time investment as well as the visible development of linguistic skills are strong advocates of the benefits of *languageing* in learners' L1, showing that CBI does not necessarily need to be built around L2 usage. However, researchers in the domain should keep in mind that this

methodology is best used for conceptualisation purposes, with the aim of targeting specific structures which are difficult to acquire naturally until later stages of formal instruction.

6.2. Limits of present research

Even though we were able to answer our research questions with the help of the analyses carried out in Chapter 5, several limits hampered the present work and should be accounted for when attempting to interpret the results.

First of all, the Limits in game design choices were already mentioned in Section 3.2.3., such as the impact of the narrative paradox for the CBI game, the difficulty of balancing tasks in a TBLT game or the influence of the peers on a supposedly individual process. Yet, results showed that notwithstanding these hurdles, learners who invested into the games efficiently were not only able to enjoy their experience, but also progress individually in a significant way. As such, it would seem that these limits were not huge obstacles in the conduct of the experiment.

A more concerning issue was that we could not gather much data on **optional verbalisation tasks** (CBI) or **final tasks** (TBLT). It would seem that making these tasks either bonus content (for the verbalisations) or somewhat vague to access through the interface (for the final task) were not appropriate. Consequently, future development using these games or others will have to account for these poor choices in design.

We also realised, after analyzing our pre-post test data, that **not all the diatheses were covered** in both tests. In fact, there was not even a single item testing the impersonal passive diathesis in either test. As a result, we could not verify the effects of the games on several specific structures, such as *Impersonal passives* and *Conversive passives* – albeit it would be difficult to give the same treatment to all pronominal constructions due to the limited

amount of items that we could include in the tests.

Yet another limitation that occurred with this type of experiment, including pre- and post-test data, was that **not all learners were able to complete both tests**. To ensure the good conduct of the experiment, all learners were supposed to attend both pre-test and post-test classes, which is not easy to control. Thus, a certain amount of data from the pre-test was not usable, even though a good amount of concerned learners had substantial playtime. This also led to lower numbers for each treatment, which perhaps could have influenced the statistical models of our analyses in a different fashion.

This limitation also brings into question the potential use of **delayed post-tests**, as we could perhaps have observed different trends in learning gains in a longitudinal study. Indeed, several CBI studies reported in the systematic review covered in Section 2.1. mentioned the importance of extending experimental designs to account for long-term effects. However, as carrying out the post-test for the target population was already difficult, we did not consider complexifying the module further.

Finally, it should be noted that the research we carried out in this thesis was, by nature, **sprawling** and might thus appear to be **unfocused**. Indeed, the work covered by our doctoral research intertwined with several disciplines:

- State-of-the-art of the way pronominal verbs are explained in textbooks used in Japanese universities (Chapter 1);
- Scoping review of Concept-based language instruction studies involving experimental research up to 2025 (Chapter 2);
- State-of-the-art of major taxonomies covering pronominal verbs and creation of a new taxonomy based on corpora analyses (Chapter 2);
- State-of-the-art of the theoretical tenets of Digital Game-based Language Learning and development of two games (Chapter 3);

- Assessment of grammatical knowledge with Forms tests design and incorporation of statistical testing through linear mixed models (Chapter 4);
- Quantitative and qualitative analysis of results to measure the development of pronominalisation (Chapter 5).

As can be seen, mastering all aspects covered in this work – descriptive linguistics, sociolinguistics, corpus linguistics, gamification & programming, assessment theories - is a challenging task, almost an overwhelming one, which could require further monitoring even after the end of this research. Yet, we believe that the research's ambitious objectives resonated with such an agenda, as by nature, concept-based studies require a deep understanding of the target concept, as well as a creative mind in order to find ways to deliver these notions. To obtain satisfying results, we needed to establish a framework including all the aforementioned aspects – as our research questions were just as ambitious. Nonetheless, we were still especially cautious when presenting each and every contribution made throughout this work, as we looked to emphasise on the strengths of our research but also on its shortcomings.

Notwithstanding the complex nature of this research and the difficulties involved, we believe that the present work offered insightful data on the development of pronominalisation with the help of mobile games, which was our main objective. To conclude this thesis, we will offer some perspectives onto the way this type of research could be improved.

6.3. Concluding notes

Throughout its relatively short history, DGLL has seen the rise of diversified ways of representing the way instructors want to convey the target knowledge. Due to the fact that mobile games usually ensure a personalised and accessible experience to players, the ever growing body of literature about their uses in education is not surprising (Hung et al., 2018). One of the most compelling aspects of games is the degree of immersion that they can bestow upon learners, who might feel less constrained in such an environment than by the traditional rules of an instructional setting (Shepherd et al., 2011). Allowing students who wish to learn abstract grammatical concepts in other ways, even if they were a minority, is a major motivation for us to keep creating educational games with a positive mindset. Yet, in order to avoid falling into the trap of making games with too little instructional scope (Sykes, 2013), the importance of founding learning tools on a sound methodology and theoretical principles should always be emphasised.

Considering the success of both CBI and TBLT games on different aspects of the conceptualisation of pronominal verbs, we are led to believe that these approaches are, in fact, **complimentary**. In that regard, there has been mention of the utility in using mixed approaches, as TBLT "*should not be seen as an alternative to more traditional form-focused approaches but can be used alongside them*" (Ellis, 2009: 221). If anything, integrating TBLT-like tasks in an environment in which concepts are embodied such as the CBI game resonates now as an appealing prospect. While the compatibility of TBLT tasks with CBI SCOBAs and verbalisations was already mentioned by authors (Fehrman, 2016, Shepherd et al., 2011), it is still a challenging project, as language exposure should be balanced to guarantee salience in

comprehension while offering enough screentime to the target knowledge.

While these games showed promising results for the specific context of Japanese learners in tertiary education, it would also be interesting to see whether this type of teaching could work with a different target audience. Most notably, as we'd seen in Section 3.1, Japanese learners usually feel comfortable with mobile devices, as practically all learners possess a mobile phone; as such, DGLL might not be as successful in other contexts in countries where phones are not as commonly spread. Implementing this type of design in different countries should thus be made with proper adaptations and adequate accessibility.

Finally, a key aspect to develop for further research is the application of CBI to other grammatical concepts. Due to its relative recency, but also the difficulty in its implementation – as the creation of thorough SCOBAs is always challenging –, CBI research has yet to uncover the effects of its methodology for many concepts. Yet, as the results in the literature as well as this thesis have suggested, CBI can be translated into various formats with overall great success on learners' conceptual acquisition. Researchers in SLA who wish to construct tools to teach grammar knowledge, in particular, are invited to consider CBI as a theoretical foundation, so as to verify its effects on learners' conceptual awareness.

Bibliography

- Acquah, E. O., & Katz, H. T. (2020). Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. *Computers & Education*, 143, 103667.
- Afifi, D. (2018). L'analyse des erreurs en production écrite chez les apprenants arabophones en classe de FLE. In *Anales de Filologia Francesa* (Vol. 26).
- Aguiló-Mora, F., & Negueruela, E. (2015). Motion-for-the-other through motion-for-the-self: The complexities of giving directions for advanced second language learners of Spanish and English. In Masuda, K. & al. (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching*, 8 (pp. 73-100). Walter de Gruyter GmbH & Co KG. Gruyter, Mouton.
- Allam, S. (2010). Les manuels de FLE se mettent-ils au goût du jour? Le cas des verbes pronominaux. In Galatanu et al. (éds.), *Enseigner les structures langagières en FLE*, Peter Lang.
- Albrecht, J. (2015). Reflexivkonstruktionen in romanischen Sprachen und germanischen Sprachen. In Lavric E. & Pöckl W. (eds.), *Comparatio delectat II, Akten der VII. Internationalen Arbeitstagung zum romanisch-deutschen und innerromanischen Sprachvergleich, Innsbruck*, 6–8(2) (2012) : 445–465.
- Aljayyousi, M. (2021). Daily Verbs: Teaching English Verb Tenses Through A Simple Video Game. *European Journal of Foreign Language Teaching*, 5(5).
- Allerton, J. (1982). *Valency and the English Verb*. London: Academic Press.
- Alsadag, A. H. (2021). Analyse linguistique des erreurs courantes en FLE chez les étudiants à l'université d'Aljufra en Libye. *Journal of Language Studies*, 4(3), 176-184.
- Amoia, M. (2011). I-FLEG: A 3D-game for learning French. In *Proceedings of the 9th International Conference on Education and Information Systems, Technologies and Applications*.
- Aoyama, R. (2020). Exploring Japanese High School Students' L1 Use in Translanguaging in the Communicative EFL Classroom. *Tesl-Ej*, 23(4).
- Apple M. T., Da Silva D. & Fellner T. (2013). *Language Learning Motivation in Japan, Multilingual Matters*, 344p.
- Armstrong, G. (2016). Towards a theory of pronominal verb constructions in Spanish. In Cuza A. et al. (eds.), *Inquiries in Hispanic Linguistics: From Theory to Empirical Evidence* : 21-38.
- Authier, J. M., & Reed, L. (1996). On the Canadian French middle. *Linguistic Inquiry*, 513-523.
- Barque, L., Candito, M., & Huyghe, R. (2018) La classification des verbes réfléchis à l'épreuve d'une annotation en corpus. *Langages*, 1-17.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, 12(1), 2.

- Beacco, J. C. (2005). *Niveau A1 pour le français: référentiel et certification (DILF) pour les premiers acquis en français*. Paris: Didier.
- Beacco, J. C., Bouquet, S., & Porquier, R. (2004). *Référentiel pour le français. Niveau B2*. Paris: Didier.
- Beinborn, L., Zesch, T., & Gurevych, I. (2014). Predicting the difficulty of language proficiency tests. *Transactions of the Association for Computational Linguistics*, 2, 517-530.
- Benhamlaoui, M., & Gánem-Gutiérrez, G. A. (2022). A Classroom Application of Concept-Based Language Instruction (C-BLI). *Language and Sociocultural Theory*, 9(1), 1-33.
- Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 827-877.
- Bibauw, S. (2022). *Learning with conversational AI: Design and effectiveness of dialogue-based computer-assisted language learning*. Doctoral dissertation. KU Leuven.
- Bień, J. (2007). Sur la nature des confusions notionnelles au sein du verbe pronominal : Diathèse. *Lublin Studies in Modern Languages and Literature*, 31(1) : 128-135.
- Bird, S. (2011). Effects of distributed practice on the acquisition of second language English syntax—ERRATUM. *Applied Psycholinguistics*, 32(2), 435-452.
- Blansitt, E. L. (1988). Datives and allatives. In Hammond M. et al. (eds.), *Studies in syntactic typology*. Amsterdam. 173-191.
- Bolker, B. M. (2015). Linear and generalized linear mixed models. *Ecological statistics: contemporary theory and application*, 309-333.
- Boons, J-P., Guillet, A. & Leclère, Ch. (1976). La structure des phrases simples en français, constructions intransitives. *Journal of Linguistics*, 15(1) : 187-192.
- Borghi, A. M., & Cimatti, F. (2010). Embodied cognition and beyond: Acting and sensing the body. *Neuropsychologia*, 48(3), 763-773.
- Bouard, B. (2011). Verbe pronominal et voix pronominale dans les grammaires françaises des 18^e et 19^e siècles. *Colloque annuel de la Société d'Histoire et d'Épistémologie des Sciences du Langage*, Paris : 733-757.
- Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14(2), 136-193.
- Brown, J. D. (2009). Statistics Corner. Questions and answers about language testing statistics: Choosing the right number of components or factors in PCA and EFA. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 13(2), 19-23.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Buescher, K., & Strauss, S. (2015). A cognitive linguistic analysis of French prepositions à, dans, and en and a sociocultural theoretical approach to

- teaching them. In Masuda, K. & al. (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching*, 8 (pp. 155-181). Walter de Gruyter GmbH & Co KG. Gruyter, Mouton.
- Burrows, C. (2008). Socio-cultural barriers facing TBL in Japan. *The Language Teacher*, 32(8), 15–19.
- Burston, J. L. (1979). The pronominal verb construction in French: An argument against the fortuitous homonymy hypothesis. *Lingua*, 48(2-3) : 147-176.
- Butler, Y. G., & Iino, M. (2005). Current Japanese reforms in English language education: The 2003 action plan. *Language Policy*, 4, 25–45.
- Butler, Y. G., Someya, Y., Fukuhara E. (2014). Online games for young learners' foreign language learning. *English Language Teaching Journal*, 68 (3), 265-275
- Butzkamm, W., & Caldwell, J. A. (2009). *The bilingual reform: A paradigm shift in foreign language teaching*. Narr Francke Attempto Verlag.
- Carpay, J. A. M. (1974). Foreign-Language Teaching and Meaningful Learning: A Soviet Russian Point of View. *ITL Review of Applied linguistics*, 25-26, 162-187.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354–380.
- Chamberlain, J., Poesio, M., & Kruschwitz, U. (2008). Phrase detectives: A web-based collaborative annotation game. In *Proceedings of the international conference on semantic systems (I-Semantics' 08)* (pp. 42-49).
- Chen, C. H. (2019). The impacts of peer competition-based science gameplay on conceptual knowledge, intrinsic motivation, and learning behavioral patterns. *Educational Technology Research and Development*, 67, 179-198.
- Chen, H. J. H., & Yang, T. Y. C. (2013). The impact of adventure video games on foreign language learning and the perceptions of learners. *Interactive learning environments*, 21(2), 129-141.
- Cheng, A., Yang, L., & Andersen, E. (2017). Teaching language and culture with a virtual reality game. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 541-549).
- Chevalier, L. (2008). Les habitudes d'enseignement-apprentissage des langues au Japon, *Seinan Gakuin University Press*, 51, 41-59.
- Chu, M.-P. & Nakamura, T. (2010). A Study of Chinese and Japanese College Students' L2 Learning Styles, *Asian Culture and History*, 2(2), 30-44.
- Cobb, T., & Horst, M. (2011). Does word coach coach words? *Calico journal*, 28(3), 639-661.
- Combettes, B. (2009). Quelle(s) description(s) grammaticale(s) pour l'enseignement ? *Repères*, 39 : 41-56.
- Comrie, B. (1985). Causative verb-formation and other verb-deriving

- morphology. In Shopen T. (ed.) *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*. Cambridge University Press : 301-348.
- Connors, K. & Ouellette, B. (1996). Describing the meanings of French pronominal-verbal constructions for students of French <--> English translation. *Language sciences*, 18(1-2) : 213-226.
- Corminboeuf, G. & Heyna, F. (2015). Nominalisation et diathèse. *Le Français moderne*, 83(1) : 34-53.
- Cornillie, F., Thorne, S. L., & Desmet, P. (2012). ReCALL special issue: Digital games for language learning: challenges and opportunities: Editorial Digital games for language learning: from hype to insight? *ReCALL*, 24(3), 243-256.
- Cornillie, F., & Desmet, P. (2016). Mini-games for language learning. In *The Routledge handbook of language learning and technology* (pp. 431-445). Routledge.
- Crain, W. (2010). *Theories of development: Concepts and applications*, 6th Ed. Upper Saddle River, Boston, MA : Prentice Hall.
- Creissels, D. (2007). Réflexivisation, transitivité et agent affecté. In Rousseau A. et al. (eds.), *L'énoncé réfléchi*, Presses Universitaires de Rennes : 83-106.
- Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organisation of information*. University of Chicago Press.
- Culioli, A. (1990). *Pour une linguistique de l'énonciation: opérations et représentations* (Vol. 1). Editions Ophrys.
- Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Houghton Mifflin Harcourt.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). Technology acceptance model. *J Manag Sci*, 35(8), 982-1003.
- DeHaan, J. (2005). Learning language through video games: A Theoretical Framework, an Evaluation of Game Genres and questions for Future Research. *SP Schaffer & ML Price* (Eds), 229-239.
- DeHaan, J., W. M. Reed, and K. Kuwanda (2010). "The effect of interactivity with a music video game on second language vocabulary recall". In: *Language Learning & Technology*, 14(2), 74-94.
- de la Fuente, A. A., & Lacroix, H. (2015). Multilingual learners and foreign language acquisition: Insights into the effects of prior linguistic knowledge. *Language Learning in Higher Education*, 5(1), 45-57.
- Delbarre, F. (2010). Nécessité de contextualisation des manuels publiés au Japon - Une étude de cas : l'expression de l'état. *Revue japonaise de didactique du français*, 5(1) : 146-166.
- Deronne, E. (2006). Des voix aux diathèses. *Article issu d'une intervention au séminaire de l'ATILF* (Nancy) de novembre 2006.
- de Schepper, K. (2010). *The space between one and two. Transitivity: Form, meaning, acquisition, and processing*, 166, 191.
- Desclés, J-P. & Guentchéva, Z. (1993). Le passif dans le système des voix du

- français, *Langages*, 109 : 73-102.
- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure: How do word-related variables and proficiency influence receptive vocabulary learning? *Language Learning*, 70(2), 349-381.
- Dickinson, P. (2010). *Implementing task-based language teaching in a Japanese EFL context* (Unpublished MA thesis). University of Birmingham, England. Published online, retrieved from the following URL: <http://www.birmingham.ac.uk/documents/college-artslaw/cels/essays/languageteaching/dickinsonltm.pdf>
- Dik, S. C. (1985). Valentie en valentie operaties in Funktionale Grammatika. Valentie in Functionele Grammatica. *Interdisciplinair Tijdschrift voor Taal-en Tekstwetenschap*, 5(2) : 95-114.
- Dixon, R. M., & Aikhenvald, A. Y. (2000). *Changing valency: Case studies in transitivity*. Cambridge: Cambridge University Press.
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, 84, 795-805.
- Dos Santos, L. M. (2020). The Discussion of Communicative Language Teaching Approach in Language Classrooms. *Journal of Education and e-Learning Research*, 7(2), 104-109.
- Duée, C. (2002). Analyse d'erreurs. Les verbes de construction pronominales dans une traduction d'un extrait de la familia Pascual Duarte, de Camilo José Cela. In *La lingüística francesa en el nuevo milenio: Universidad de Lleida*, 8, 9 y 10 de noviembre de 2001 (pp. 221-228). Milenio.
- Ebrahimzadeh, M., & Alavi, S. (2016). Motivating EFL students: E-learning enjoyment as a predictor of vocabulary learning through digital video games. *Cogent Education*, 3(1), 1255400.
- Edelman, G. M. (2004). *Wider than the sky: The phenomenal gift of consciousness*. Yale University Press: New Haven, Connecticut.
- Ellis, R. (2000). Task-based research and language pedagogy. *Language teaching research*, 4(3), 193-220.
- Ellis, R. (2003). Designing a task-based syllabus. *RELC Journal*, 34(1), 64-81.
- Ellis, R. (2014). Taking the critics to task: The case for task-based teaching. In: *Proceedings of CLASIC*, pp. 103-117.
- Erdocia, K., & Laka, I. (2018). Negative transfer effects on L2 word order processing. *Frontiers in Psychology*, 9, 337.
- Evrard, I. J. M. S. (2003). La diathèse en français : un essai de synthèse. *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*: 303-316.
- Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *Relc Journal*, 31(2), 105-119.

- Fazilatfar, A., Jabbari, A. & Harsij, R. (2017). Concept-based Instruction and Teaching English Tense and Aspect to Iranian School Learners. *Issues in Language Teaching*, 6(1), 145-179.
- Fehrmann, I. (2016). Teaching the form-function mapping of German 'prefield' elements using Concept-Based Instruction. *Yearbook of the German Cognitive Linguistics Association*, 4(1), 153-170.
- Fellbaum, C., & Zribi-Hertz, A. (1989). La construction moyenne en français et en anglais : étude de syntaxe et de sémantique comparées. *Recherches linguistiques de Vincennes*, 18, 19-57.
- Ferreira, M. (2005). *An application of the concept-based approach to academic writing instruction*. Ph.D. diss., Pennsylvania State University, University Park, PA.
- Fogal, G. G. (2016). *Pedagogical Stylistics and Concept-Based Instruction: An Investigation into the Development of Voice in the Academic Writing of Japanese University Students of English*. Ph.D. diss., University of Toronto (Canada).
- François, J. (2011). Quelles sont les origines des verbes essentiellement pronominaux du français ? In Dessi Schmidt S. et al. (eds.), *Rahmen des Sprechens : Beiträge zur Valenztheorie, Varietätenlinguistik, Kreolistik, Kognitiver und Historischer Semantik*, Peter Koch, 60 : 33-42.
- François, T., Gala, N., Watrin, P., & Fairon, C. (2014, May). FLELex: a graded lexical resource for French foreign learners. In *International conference on Language Resources and Evaluation (LREC 2014)*.
- Frazier, E. G. G. (2013). *Concept-based teaching and Spanish modality in heritage language learners: A Vygotskian approach*. Ph.D. diss., University of Massachusetts Amherst.
- Gaatone, D. (1975). Réflexions sur les verbes pronominaux réfléchis et réciproques. *Folia Linguistica*, 8(1-4): 199-222.
- Gal'perin, P. I. (1989). Organisation of mental activity and the effectiveness of learning. *Soviet Psychology*, 27 (3): 65-82.
- Gal'perin, P. I. (1992). Stage-by-stage formation as a method of psychological investigation. *Journal of Russian and East European Psychology*, 30 (July/August): 60-80.
- Galton, F. (1886) Regression towards mediocrity in hereditary stature, *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246-263.
- Gánem-Gutiérrez, G. A. & Harun, H., (2011). Verbalisation as a mediational tool for understanding tense-aspect marking in English: an application of Concept-Based Instruction. *Language Awareness*, 20 (2), 99-119
- Garcia, P. N. (2012). *Verbalizing in the second language classroom: The development of the grammatical concept of aspect*. Ph.D. diss., University of Massachusetts Amherst.

- García, P. N. (2017). Implementing concept-based instruction in the heritage language classroom: A pedagogical proposal. *EuroAmerican Journal of Applied Linguistics and Languages*, 4(1), 1-19.
- Gee, J. P. (2003). From video games, learning about learning. *Chronicle of Higher Education*, 49(41 SEC. 2), B13.
- Gee, J. P. (2007). *Good video games + good learning: Collected essays on video games, learning, and literacy*. Peter Lang
- Geniušienė, E., (1987). *The typology of reflexives*. Berlin: Mouton de Gruyter.
- Godwin, K. E., Seltman, H., Almeda, M., Davis Skerbetz, M., Kai, S., Baker, R. S., & Fisher, A. V. (2021). The elusive relationship between time on-task and learning: Not simply an issue of measurement. *Educational Psychology*, 41(4), 502-519.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- González, J. M. M., de Castro Castro, C., & Millán, A. I. B. (2021). The adventure of an adventure videogame in French. *Digital Education Review*, (39), 257-269.
- Gorsuch, G. J. (1998). Yakudoku EFL instruction in two Japanese high school classrooms: An exploratory study. *JALT journal*, 20(1), 8.
- Göbel, S., de Carvalho Rodrigues, A., Mehm, F., & Steinmetz, R. (2009). Narrative game-based learning objects for story-based digital educational games. *Narrative*, 14, 16.
- Graham, S., & Taylor, A. Z. (2016). Attribution theory and motivation in school. *Handbook of motivation at school* (pp. 11-33). Routledge.
- Griggs, P. (2007). *Perspective sociocognitive sur l'apprentissage des langues étrangères* (p. 208). L'Harmattan.
- Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA, MIT Press.
- Gyllstad, H., Vilkaitė, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL-International Journal of Applied Linguistics*, 166(2), 278-306.
- Halbhuber, D., Thomaschke, R., Henze, N., Wolff, C., Probst, K., & Bogon, J. (2023). Play with my Expectations: Players Implicitly Anticipate Game Events Based on In-Game Time-Event Correlations. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (pp. 386-397).
- Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language teaching*, 45(3), 271-308.
- Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins & M. Swain (Eds.), *The development of second language proficiency* (pp. 7-25). Cambridge: Cambridge University Press.
- Harun, H., Massari, N., & Behak, F. P. (2014). Use of L1 as a mediational tool for understanding tense/aspect marking in English: An application of Concept-Based Instruction. *Procedia-Social and Behavioral Sciences*,

- 134, 134-139.
- Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1) : 1-33.
- Heidinger, S. (2015). Causalness and the encoding of the causative–anticausative alternation in French and Spanish¹. *Journal of Linguistics*, 51(3), 562-594.
- Heidinger, S., & Schäfer, F. (2008). The French Reflexive Passive and Anticausative. In Fagard B. et al. (eds), *Evolutions en français. Etude de linguistique diachronique*, Berlin, Peter Lang : 135-152.
- Hess, E. K., & Amory, M. (2022). Promoting Korean learners' conceptual development of honorifics through concept-based language instruction. *Language Teaching Research*, 13621688221135567.
- Hirashima, R. (2010). 教科書分析から見たフランス語教授法の改善策について. *Revue japonaise de didactique du français*, 5(1), 236-248.
- Hitosugi, C. I., Schmidt, M., & Hayashi, K. (2014). Digital game-based learning (DGBL) in the L2 classroom: The impact of the UN's off-the-shelf videogame, Food Force, on learner affect and vocabulary retention. *Calico Journal*, 31(1), 19-39.
- Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C. & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, 126, pp. 89–104.
- Hwang, G. J., Hsu, T. C., Lai, C. L., & Hsueh, C. J. (2017). Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns. *Computers & Education*, 106, 26-42.
- Iaremenko, N. V. (2017). Enhancing English language learners' motivation through online games. *Інформаційні технології і засоби навчання*, (59, вип. 3), 126-133.
- Irie, K. (2003). What do We know about the Language Learning Motivation of University Students in Japan ? Some Patterns in Survey Studies. *Japan Association for Language Teaching (JALT)*, 25 n°1, 86-101.
- Ishikawa, F. (2011). Le FLE comme matière universitaire en contexte « hétéroglotte » et extra-européen : statut académique et représentations sociales du FLE au Japon. *Le Français sur Objectifs Universitaires*, 301-312.
- Jackendoff, R. S. (1992). *Semantic structures* (Vol. 18). MIT press.
- Jarvis, S. (2011). Conceptual transfer: Crosslinguistic effects in categorization and construal. *Bilingualism: Language and cognition*, 14(1), 1-8.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language learning*, 64(1), 160-212.
- Joffre, M. D. (2018). La synonymie partielle entre deux signifiants : un moteur d'évolution et de création. Le cas de la genèse, en latin, du pronominal français. *Orbis Linguarum*, 50 : 103-114.

- Jones, L. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning & Technology*, 8(3), 122-143.
- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4), 396-403.
- Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., & Rich, B. L. (2007). Self-efficacy and work-related performance: the integral role of individual differences. *Journal of applied psychology*, 92(1), 107.
- Juul, J. (2005). Games telling stories. *Handbook of computer game studies*, 219-226.
- Kailuwet, R. (2007). La classe P dans Les Verbes français et les verbes de sentiment. *Langue française*, 1: 33-39.
- Kallulli, D. (2006). A unified analysis of passives and anticausatives. *Proceedings of Sinn und Bedeutung*, 10(1) : 171-182.
- Kambaru, R. (2023). Pedagogical Implications of Concept-Based Language Instruction for English Education in Japan: A Literature Review of Concept-Based Language Instruction Research. *東京大学大学院教育学研究科紀要*, 62, 65-73.
- Karan, D. (2016). Analyse des erreurs syntaxiques des productions écrites chez les apprenants Turcs dans l'enseignement du FLE suivants l'approche communicative. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 29(1).
- Karpov, Y. V. (2003). Vygotsky's doctrine of scientific concepts: Its role for contemporary education. *Vygotsky's educational theory in cultural context*, 65-82.
- Karpova, S. N. (1977). The realisation of the verbal composition of speech by preschool children. *With a forew. for the Engl.* ed. by Luria, A. R. The Hague, Mouton.
- Karasimos, A., & Zorbas, V. (2019) From multi-genres computer games to multi-skills language teaching and learning: A theoretical proposal. In *Book of Proceedings* (p. 292).
- Kessler C. & Siary G. (2008). *France-Japon : histoire d'une relation inégale*, texte issu de La vie des idées (laviedesidees.fr), 1-12.
- Kikuchi, K., & Sakai, H. (2009) Japanese learners' demotivation to study English: A survey study. *JALT Journal*, 31, 183-204.
- Kim, J. (2013). *Developing conceptual understanding of sarcasm in a second language through concept-based instruction*. Unpublished doctoral dissertation. Pennsylvania State University, University Park, PA.
- Kim, E. S., & Willson, V. L. (2010). Evaluating pre-test effects in pre-post studies. *Educational and psychological measurement*, 70(5), 744-759.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*.
- Knekta, E., & Eklöf, H. (2015). Modeling the test-taking motivation construct through investigation of psychometric properties of an expectancy-value-based questionnaire. *Journal of Psychoeducational Assessment*, 33(7),

662-673.

- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational researcher*, 49(4), 241-253.
- Kraus, V. (1997). *L'apprentissage du français langue seconde chez les immigrants chinois, coréens et japonais à Montréal*. UWSpace. Retrieved from <http://hdl.handle.net/10012/14293>
- Kulikov, L. (2011). Voice typology. In Song J. J. (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford University Press : 368–398.
- Labelle, M., & Doron, E. (2010). Anticausative derivations (and other valency alternations). *French. Probus*, 22(2) : 303-316.
- Lai, W. (2012). *Concept-based foreign language pedagogy: Teaching the Chinese temporal system*. Ph.D. diss., Pennsylvania State University, University Park, PA.
- Lapkin, S., & Swain, M. (2000). Task outcomes: A focus on immersion students' use of pronominal verbs in their writing. *Canadian Journal of Applied Linguistics*, 3(1-2), 7-22.
- Lantolf, J. P. (2011). Integrating sociocultural theory and cognitive linguistics in the second language classroom. In Hinkel, E. (Ed.), *Handbook of research in second language teaching and learning*, 2 (pp. 303-318). Routledge.
- Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural theory and second language development. In Van Patten, B. & al. (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 207-226). Taylor and Francis.
- Lee, H. (2012). *Concept-based approach to second language teaching and learning: Cognitive linguistics-inspired instruction of English phrasal verbs*. Unpublished PhD. dissertation. Pennsylvania State University.
- Lauwers, P., & Tobback, E. (2013). Émotions, subjectivité et morphosyntaxe : l'impact de la clôture actancielle sur les verbes pronominaux à attribut de l'objet. *Langue française*, 4 : 47-64.
- Lee, H. (2016). Concept-Based Instruction: Imagistic and Metaphorical Understanding of Phrasal Verbs. *English Teaching*, 71(4), 167-191.
- Lee, J. (2019). Time-on-task as a measure of cognitive load in TBLT. *Journal of Asia TEFL*, 16(3), 958.
- Le Querler, N. (2012). Valence et complémentation : L'exemple des verbes trivalents en français contemporain. *Annales de Normandie*, 62(2) : 175-188.
- Lewis, M. W. (2007). Analysis of the roles of “serious games” in helping teach health-related knowledge and skills and in changing behavior. *Journal of Diabetes Science and Technology*, 1(6), 918-920.
- Li, S., Ellis, R., & Zhu, Y. (2016). Task-based versus task-supported language instruction: An experimental study. *Annual Review of Applied Linguistics*, 36, 205-229.

- Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, 65(S1), 185-207.
- Lindstromberg, S. (2016). Inferential statistics in Language Teaching Research: A review and ways forward. *Language Teaching Research*, 20 (6), 741-768.
- Liu, S., & Hsieh, C. Y. C. (2020). Developing metaphorical awareness and competence in Chinese as a foreign language through concept-based instruction. *Foreign Language Annals*, 53(3), 478-504.
- Louchart, S., & Aylett, R. (2003, September). Solving the narrative paradox in VEs—lessons from RPGs. In *International workshop on intelligent virtual agents* (pp. 244-248). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Lyster, R. (1999). Immersion. In B. Spolsky (Ed.), *Concise encyclopedia of educational Linguistics* (pp. 626-632). Oxford: Pergamon.
- Lyster, R. (2004). Research on form-focused instruction in immersion classroom: Implications for theory and practice. *Journal of French language studies*, 14(3), 321-341.
- Mackey, A. Goo, J. (2007). *Interaction research in SLA: A meta-analysis and research synthesis*.
- Marsden, E., & Torgerson, C. J. (2012). Single group, pre-and post-test research designs: Some methodological concerns. *Oxford Review of Education*, 38(5), 583-616.
- Masuda, K., & Labarca, A. (2015). Schematic diagram use and languaging quality in learning Japanese polysemous particles ni and de. In Masuda, K. & al. (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching*, 8 (pp. 203-232). Walter de Gruyter GmbH & Co KG.
- Mayaux, C. (2007). France-Japon : regards croisés. *Littératures de langue française*, VII, 39-52.
- Melis, L. (1990a). Pronominal Verbs in Old and Modern French, or How prototypes can be restructured on the basis of permanent meaning effects. *Belgian Journal of Linguistics*, 5(1) : 87-108.
- Melis, L. (1990b). *La voie pronominale. La systématique des tours pronominaux en français moderne*, Duculot, Champs linguistiques, 157p.
- Midway, S. (2022). *Data analysis in R*. Retrieved April, 7, 2024.
- Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: the link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4(1), 151-172.
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), 1-24.
- Mondria, J., & Wiersma, B. (2004). Receptive, productive and receptive+productive L2 vocabulary learning: What difference does it make. *Vocabulary in a second language: Selection, acquisition and testing*, 79-

100.

- Nakamura K. (1996). Du vocabulaire à la grammaire. Quelques verbes difficiles à utiliser pour les apprenants japonais. *Bulletin du Centre Pédagogique l'Université Otemon*, 15, 65-71.
- Nakamura K. (2004). *Problèmes typiques des apprenants japonophones du français*. Tokyo : Surugadai-Shuppansha, 135p.
- Nakamura, T. (2014). Sur les constructions de la forme SN se faire ADV. *Langages*, 2 : 95-106.
- Nakamura, S., Phung, L., & Reinders, H. (2021). The effect of learner choice on L2 task engagement. *Studies in Second Language Acquisition*, 43(2), 428-441.
- Negueruela, E. (2003). *Systemic-theoretical instruction and L2 development: A sociocultural approach to teaching-learning and researching L2 learning*. Ph.D. diss., Pennsylvania State University, University Park, PA.
- Negueruela, E., & Lantolf, J. P. (2006). Concept-based instruction and the acquisition of L2 Spanish. In Salaberry, R., & Lafford, B. A. (Eds.), *The Art of Teaching Spanish: Second Language Acquisition from Research to Praxis* (pp. 79-102). Georgetown University Press.
- Negueruela, E. (2008). Revolutionary pedagogies: learning that leads development in the second language classroom. In James P. Lantolf & Matthew Poehner (eds.), *Sociocultural theory and the teaching of second languages*, 189–227. London: Equinox Publishers.
- Negueruela, E. (2013). Metalinguistic knowledge becoming intrapsychological tools. Concepts as tools for second language development. In Roehr, K., & Gánem-Gutiérrez, G. A. (Eds.), *The metalinguistic dimension in instructed second language learning* (pp. 221-242). A&C Black.
- Nguyen, V. N. (2004). *Pronoms personnels du français et du vietnamien: une étude contrastive et analyse des erreurs en FLE*. Ph.D. diss., University of Montreal.
- Nishino, T. (2008). Japanese secondary school teachers' beliefs and practices regarding communicative language teaching: An exploratory survey. *JALT Journal*, 30(1), 27.
- Noda, M., & O'Regan, J. P. (2020). L1 marginalisation in Japan: Monolingual instrumentalism and the discursive shift against yakudoku in the Japanese government's Course of Study. *Current issues in language planning*, 21(2), 135-152.
- Noreillie, A.-S., Kestemont, B., Heylen, K., Desmet, P., & Peters, E. (2018). Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages. *ITL - International Journal of Applied Linguistics*, 169(1), 212–231.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, 50(3), 417-528.
- Nowlan, A., & Samuëll, C. (2020). Communicative language teaching:

- Contemporary applications in Japan. *Kwansei Gakuin University Humanities Review*, 24, 77-87.
- Nuryakin, N., Rakotoarizaka, N. L. P., & Musa, H. G. (2023). The effect of perceived usefulness and perceived easy to use on student satisfaction the mediating role of attitude to use online learning. *APMBA (Asia Pacific Management and Business Application)*, 11(3), 323-336.
- O'Donnell, K. (2003). Uncovering first year students' language learning experiences, attitudes, and motivations in a context of change at the tertiary level of education. *JALT Journal*, 25(1), 31-62.
- Ohta, A. S., & Tsujihara, R. (2017). Teaching addressee honorifics in Japanese using concept-based instruction (CBI): A qualitative, intact classroom study. In *Conference of the American Association for Applied Linguistics* (pp. 18-21), Portland, Oregon.
- Ord, C. A. (2014). *The marking of aspect in the past tense in French: The role of languaging abstract grammatical concepts*. Ph.D. diss., School of Languages and Comp Cultural Studies, University of Queensland.
- Ortega, L. (2009). Studying writing across EFL contexts: Looking back and moving forward. *Writing in foreign language contexts: Learning, teaching, and research*, 232-255.
- Parry, M. (1998). The reinterpretation of the reflexive in Piedmontese: impersonal SE constructions. *Transactions of the Philological Society*, 96(1): 63-116.
- Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition*, 2(3), 209-230.
- Perez-Beltrachini, L., Gardent, C., & Kruszewski, G. (2012). Generating grammar exercises. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP* (pp. 147-156).
- Pérez-Tattam, R., Gathercole, V. C. M., Yavas, F., Stadthagen-González, H., & Anrrich, G. (2013). Measuring grammatical knowledge and abilities in bilinguals: Implications for assessment and testing. *Issues in the assessment of bilinguals*, 111-129.
- Petersilka, C. (2012). Une approche syntaxique et lexicale des verbes pronominaux. *Studia Romanica Posnaniensia*, 39(3), 33-48.
- Peterson, M. (2010). Massively multiplayer online role-playing games as arenas for second language learning. *Computer assisted language learning*, 23(5), 429-439.
- Philippette, T. (2014). Gamification: Rethinking 'playing the game' with Jacques Henriot. In *Rethinking gamification* (pp. 187-200). Meson.
- Pienkoss, L. (2021). *Mediating the German Case System Through Concept-Based Instruction in an Online Learning Environment*. Minnesota State University, Mankato.
- Pintard, A., & François, T. (2020). Combining expert knowledge with frequency information to infer CEFR levels for words. In *Proceedings of the 1st Workshop on Tools and Resources to Empower People with REading*

- Difficulties* (READI) (pp. 85-92).
- Pinter, A. (2005). Task repetition with 10-year-old children. In *Teachers exploring tasks in English language teaching* (pp. 113-126). London: Palgrave Macmillan UK.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of educational psychology*, 90(1), 25.
- Polizzi, M. C. (2013). *The development of Spanish aspect in the second language classroom: Concept-based pedagogy and dynamic assessment*. Ph.D. diss., University of Massachusetts Amherst.
- Poole, E. (2014). Deconstructing quirky subjects. *Proceedings of the 45th Meeting of the North East Linguistic Society* (NELS 45), Amherst : 1-14.
- Primus, B. (2006). Mismatches in semantic-role hierarchies and the dimensions of role semantics. In Bornkessel I. et al. (eds.), *Semantic Role Universals and Argument Linking: Theoretical, Typological, and Psycholinguistic Perspectives*, Mouton : 53-88.
- Purpura, J. E. (2013). Assessment of grammar. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics: Assessment and Evaluation* (pp. 1-10). Oxford, UK: Wiley.
- Purpura, J. E. (2014). Cognition and language assessment. In A. J. Kunnan (Ed.), *Companion to Language Assessment* (pp. 1452-1476). Oxford, UK: Wiley-Blackwell.
- Qin, L., Ouyang, X., & Ren, W. (2023). Decoding EFL learners' internalization of English modal verbs in concept-based language instruction: Integrating SCOBAs with languaging. *The Modern Language Journal*, 107(S1), 24-45.
- Reinhardt, J., & Thorne, S. L. (2020). Digital Games as Language-Learning Environments. *Handbook of game-based learning*, 409.
- Reitz, L., Sohny, A., Lochmann, G. (2016). VR-based gamification of communication training and oral examination in a second language. *International Journal of Game-Based Learning*, 6(2), 391-404.
- Renoud, L. (2016). *Vers une approche par concepts pour l'apprentissage des temps du passé en français langue étrangère dans le contexte de l'université au Japon*. Ph.D. diss., Université de Lyon.
- Richards, J. C. (2006). *Communicative language teaching today*. Singapore: SEAMEO Regional Language Centre.
- Rimmer, W. (2006). Measuring grammatical complexity: The Gordian knot. *Language testing*, 23(4), 497-519.
- Rivière, N. (1997). Le pronominal, les rôles actanciels et la diathèse. *Cahiers Charles V*, 23(1) : 7-46.
- Rivière, N. (2018). La pronominalisation et les auxiliaires de modalité. In François J. et al. (eds.), *Morphosyntaxe et sémantique du verbe relations actancielles, voix, aspect et statut grammatical en français et en arabe*, Université Caen : 60-80.

- Robinson, P., & Ellis, N. C. (2008). *Handbook of cognitive linguistics and second language acquisition* (Vol. 270). London: Routledge.
- Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science*, 16, 183–186.
- Rousseau, A. (2014). Propositions pour une description ordonnée des « voix » et des « diathèses »: problématique, statut et conceptualisation du «moyen». *Langages*, 2 : 21-34.
- Roussel F. (2006). Mieux faire entrer « l'altérité » dans la classe. *Le FLE au Japon : un regard croisé franco-japonais*, VII, 35-41.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1), 54-67.
- Sabbah, S. (2015). Negative transfer: Arabic language interference to learning English. *Arab world English journal (AWEJ) Special issue on Translation*, (4).
- Saito, S. (2020). When is it most appropriate to control for initial scores? A comparison of examination methods for two-wave panel survey data changes. *Tutorials in Quantitative Methods for Psychology*, 16(5), 457-466.
- Salaberry, M. R. (2001). The use of technology for second language learning and teaching: A retrospective. *The modern language journal*, 85(1), 39-56.
- Salen, K., & Zimmerman, E. (2003). This is not a game: Play in cultural environments. In *Proceedings of DiGRA 2003 Conference: Level Up*.
- Sandberg, J., Maris, M., Hoogendoorn, P. (2014). The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education*, 76, 119-130.
- Schäfer, F., & Vivanco, M. (2016). Anticausatives are weak scalar expressions, not reflexive expressions. *Glossa: a journal of general linguistics*, 1(1) : 1-18.
- Schmitz, B., R. Klemke, and M. Specht (2012). Effects of mobile gaming patterns on learning outcomes: a literature review. *International Journal of Technology Enhanced Learning*, 4(5-6), 345–358.
- Schmoll, L. (2016). L'emploi des jeux dans l'enseignement des langues étrangères: du traditionnel au numérique. *Sciences du jeu*, vol. 5.
- Schøsler, L. (1999). Réflexions sur l'optionnalité des compléments d'objet direct en latin, en ancien français, en moyen français et en français moderne. In *La langue, les signes et les Êtres : actes du colloque de l'Institut d'Etudes Romanes de l'Université de Copenhague* : 9-28
- Serrano-Lopez, M. & Poehner, M. E. (2008). Materializing linguistic concepts through 3-D clay modeling: A tool-and-result approach to mediating L2 Spanish development. In J. P. Lantolf, & M. E. Poehner (Eds.). *Sociocultural Theory and the Teaching of Second Languages* (pp.321-346). London, U.K.: Equinox.

- Shepherd, J. J., Doe, R. J., Arnold, M., Cheek, N., Zhu, Y., & Tang, J. (2011). Lost in the middle kingdom: a second language acquisition video game. In *Proceedings of the 49th Annual Southeast Regional Conference*, 290-294.
- Shibatani, M. (2006). On the conceptual framework for voice phenomena. *Linguistics*, 44(2) : 217-269.
- Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research*, 16(2), 253-279.
- Sijtsma, K., & Pfadt, J. M. (2021). Part II: On the use, the misuse, and the very limited usefulness of Cronbach's alpha: Discussing lower bounds and correlated errors. *Psychometrika*, 86(4), 843-860.
- Skehan, P. (1996). Second Language Acquisition Research and Task-Based Instruction'in J. Willis, and D. Willis (eds.) *Challenge and Change in Language Teaching*. Oxford: Heinemann, 17, 30.
- Song, F., Hooper, L. & Loke, Y.K. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, 5(1): 51-81
- Soyoof, A. (2018). Video games: A springboard for enhancing students' L2C. *International Journal of Pedagogies and Learning*, 13(2), 137-151.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for research in mathematics education*, 36(5), 404-411.
- Steel, A. (2020). *LEVEL UP!: Japanese second language acquisition in the classroom through video games*. University of Washington.
- Stroup, W. W. (2013). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. CRC Press.
- Su, F. et al. (2021). "A comparative review of mobile and non-mobile games for language learning". In: SAGE Open, 11(4), pp. 1–20.
- Suñer, F. & Roche, J. (2019). Embodiment in concept-based L2 grammar teaching: The case of German light verb constructions. *Journal of Cognitive Science* 20(3): 365-399
- Suzuki E. (2001). La grammaire dans l'enseignement/apprentissage universitaire du français langue étrangère au Japon. *Ela, Études de linguistique appliquée*, 2 (n°122), 143-151.
- Suzuki, E. (2003). Confrontation des cultures d'enseignement et d'apprentissage dans la classe de japonais langue étrangère en France et de français langue étrangère au Japon dans l'enseignement supérieur. In *Proceedings of the 1er Colloque international de l'Association Des Centres Universitaires d'Études Françaises pour Étrangers*.
- Swain, M. (1997). French immersion programs in Canada. In J. Cummins & D. Corson (Eds.), *Encyclopedia of language and education*, vol. 5: bilingual education. (pp. 260.299) Amsterdam: Kluwer Academic Publishers.
- Swain, M., Lapkin, S., Knouzi, I., Suzuki, W., & Brooks, L. (2009). *Languageing: University students learn the grammatical concept of voice*

- in French. *The Modern Language Journal*, 93(1), 5-29.
- Swain, M., & Lapkin, S. (2013). A Vygotskian sociocultural perspective on immersion education: The L1/L2 debate. *Journal of immersion and content-based language education*, 1(1), 101-129.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and instruction*, 12(3), 185-233.
- Sykes, J. M. (2013). 'Just' playing games? A look at the use of digital games for language learning. *The Language Educator*, 8(5), pp. 32-35.
- Tahar, A., Riyadh, H. A., Sofyani, H., & Purnomo, W. E. (2020). Perceived ease of use, perceived usefulness, perceived security and intention to use e-filing: The role of technology readiness. *The Journal of Asian Finance, Economics and Business*, 7(9), 537-547.
- Tang, J., & Wang, S. (2024). A Review of Concept-based Language Instruction and Its Application in Foreign Language Teaching-learning Process. *Lecture Notes in Education Psychology and Public Media*, 58, 32-43.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53.
- Theodorsen, J. H. (2015). *L2 acquisition from video games with minimal exposure* (Master's thesis, NTNU).
- Thorne, S. L., Fischer, I., & Lu, X. (2012). The semiotic ecology and linguistic complexity of an online game world. *ReCALL*, 24(3), 279-301.
- Thorne, S., J. Reinhardt, & P. Golombek. (2008) Mediation as objectification in the development of professional academic discourse: A corpus-informed curricular innovation. In J. P. Lantolf & M. Poehner (Eds.), *Sociocultural Theory and the Teaching of Second Languages* (pp. 256-284). London: Equinox.
- Tiwari, V., Tiwari, P. S. N., & Sharma, K. (2014). Academic motivation and school performance among students. *Indian Journal of Health and Wellbeing*, 5(4), 437A.
- Twyman, T., Ketterlin-Geller, L., McCoy, J. D., & Tindal, G. (2003). Effects of concept-based instruction on an English language learner in a rural school: A descriptive case study. *Bilingual Research Journal*, 27(2), 259-274.
- Uner, O., Tekin, E., & Roediger III, H. L. (2022). True-false tests enhance retention relative to rereading. *Journal of Experimental Psychology: Applied*, 28(1), 114.
- Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest-posttest control group design. *Structural equation modeling: a multidisciplinary journal*, 24(3), 428-450.
- Van Compernelle, R. A. (2012). *Developing sociopragmatic capacity in a second language through concept-based instruction*. Unpublished doctoral dissertation. Pennsylvania State University, University Park, PA.
- Van Gog, T. (2013). Time on task. *International guide to student achievement*

- (pp. 432-433). Routledge.
- Van Trijp, R. (2015). Cognitive vs. generative construction grammar: The case of coercion and argument structure. *Cognitive Linguistics*, 26(4) : 613-632.
- Vegel, A. (2018). Critical perspective on language learning: TBLT and digital games. In: *Proceedings of the TBLT in Asia 2018 Conference*, pp. 113-124.
- Vincent, F., & Lefrançois, P. (2013). L'opposition inductif/déductif en enseignement de la grammaire: un débat à nuancer. *Revue des sciences de l'éducation*, 39(3), 471-490.
- Vygotsky, L. S. (1987). Thinking and Speech. In R. W. Rieber, & A. S. Carton (Eds.), *The Collected Works of L. S. Vygotsky (Vol. 1), Problems of General Psychology* (pp. 39-285). New York: Plenum Press. (Original Work Published 1934)
- Wang, T. (2023). *The concept of conceptual metaphor as a psychological tool for L2 development: A conceptual metaphor theory-informed concept-based language instruction*. The Pennsylvania State University.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27, 33-52.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second language acquisition*, 30(1), 79-95.
- Weidenbusch, W. (2016). 11 La diathèse verbale (voix active, passive, pronominale etc.) comme problème de traduction. *Manuel de traductologie*, 5 : 231-255.
- White, B. J. (2012). A conceptual approach to the instruction of phrasal verbs. *The Modern Language Journal*, 96(3), 419-438.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary educational psychology*, 25(1), 68-81.
- Willis, J. (1996). A flexible framework for task-based learning. *Challenge and change in language teaching*, 52, 62.
- Wilkens, R., Pintard, A., Alfter, D., Folny, V., & François, T. (2023, December). TCFLE-8: a Corpus of Learner Written Productions for French as a Foreign Language and its Application to Automated Essay Scoring. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 3447-3465).
- Williams, L., Abraham, L. B., & Negueruela, E. (2013). Using concept-based instruction in the L2 classroom: Perspectives from current and future language teachers. *Language Teaching Research*, 17(3), 363-381.
- Williams, L. (2016). Authenticity and Pedagogical Grammar: A Concept-based Approach to Teaching French Auxiliary Verbs. In Van Compernelle, R. A., & McGregor, J. (Eds.), *Authenticity, language, and interaction in second language contexts* (pp. 35-60), Multilingual Matters.
- Yancey, K., Pintard, A. & Francois, T. (2021). Investigating readability of

- French as a foreign language with deep learning and cognitive and pedagogical features. *Lingue e linguaggio*, 20(2), 229-258.
- Yanez-Prieto, M. D. C. (2008). *On literature and the secret art of (im)possible worlds: Teaching literature-through-language*. Ph.D. diss., Pennsylvania State University, University Park, PA.
- Yang, Y., Lei, Z., Zhang, S., Shi, H., & Li, S. (2016). Metric embedded discriminative vocabulary learning for high-level person representation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 30, No. 1).
- Zhang, X. (2014). *The teachability hypothesis and concept-based instruction: Topicalisation in Chinese as a second language*. Unpublished doctoral dissertation. Pennsylvania State University, University Park, PA.
- Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal*, 93(4), 489-511.
- Ziegler, N. (2016). Taking technology to task: Technology-mediated TBLT, performance, and production. *Annual Review of Applied Linguistics*, 36, 136-163.
- Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going?. *Computer Assisted Language Learning*, 34(5-6), 751-777.
- Zribi-Hertz A. (1982). La construction du français et son statut dans le triangle : moyen-passif-réfléchi. *Lingvisticæ Investigationes*, 6(2): 345-401.
- Zribi-Hertz, A. (1987). La réflexivité ergative en français moderne. *Le Français moderne*, 55(1/2) : 23-54.
- Zúñiga, F., & Kittilä, S. (2010). *Benefactives and malefactives: Typological perspectives and case studies*. (Vol. 92). John Benjamins Publishing.

Appendix

Annex 1: Comparison between taxonomies of pronominal verbs

Annex 1: Comparison of the classifications of the sample verbs

Verb	Zribi-Hertz (1987)	Boons et al. (1976)	Geniušienė (1987)	Desclés et al. (1993)	Bien (2007)	Petersilka (2012)	Barque et al. (2018)	Our typology
<i>Se laver</i>	Reflexive	Reflexive Body-part	Reflexive	Reflexive	Active	Attribute-value	Reflexive Autobenef active	Canonical refl. Reflexive ben. Conversive pass.
<i>Se regarder</i>	Reflexive	Reflexive Reciprocal	Reflexive	Reflexive Reciprocal	Active	Attribute-value Reciprocal	Reciprocal	Canonical refl. Agentless pass. Canonical recip.
<i>Se raser</i>	Reflexive	Reflexive Body-part	Reflexive	Reflexive	Active	Attribute-value Reciprocal	Reflexive Autobenef active	Canonical refl. Reflexive ben.
<i>S'aimer</i>	Reciprocal	Reciprocal	Reflexive	Reciprocal	Active	Reciprocal	Reciprocal	Canonical recip.
<i>Se voir</i>	Reciprocal	Reciprocal	Reflexive	Reciprocal	Active	Reciprocal Mediopassive	Reciprocal Autonomo us (construction)	Canonical recip. Agentless pass.
<i>Se parler</i>	Reciprocal	Reciprocal	Reflexive	Reciprocal	Active	Reciprocal Mediopassive	Reciprocal Mediopassive	Canonical refl. Agentless pass.
<i>Se boire</i>	Passive	Phantom agent	Quasi-passive	Passive	Passive	Mediopassive	Mediopassive	Agentless pass. Reflexive ben.
<i>Se vendre</i>	Passive	Phantom agent	Quasi-passive	Passive	Passive	Mediopassive	Mediopassive	Agentless pass. Canonical refl.
<i>S'utiliser</i>	Passive	Phantom agent	Quasi-passive	Passive	Passive	Mediopassive	Mediopassive	Agentless pass.
<i>S'évanouir</i>	Subjective	Intransitive	Pseudo-reflexive	Medio-passive	Medio-active	Reflexiva tantum	Autonomo us (syntax)	Anticausative
<i>Se repentir</i>	Subjective	Possessive reflexive	Deaccusative	Middle	Medio-active	Reflexiva tantum	Autocausative	Autocausative
<i>Se douter</i>	Subjective	Possessive reflexive	Deaccusative	Middle	Medio-active	Completive	Parallel	Autocausative
<i>Se jouer</i>	Subjective	Possessive reflexive	Deaccusative	Middle	Medio-active	Completive	Mediopassive	Agentless pass.

							Parallel	Autocausative
<i>Se mourir</i>	Subjective	Intransitive	Autocausative	Medio-passive	Medio-passive	Reflexivatum	Autonomous (syntax)	Autocausative
<i>Se casser</i>	Subjective	Intransitive	Decausative	Medio-passive	Medio-passive	Anticausative	Anticausative	Anticausative Reflexive ben.
<i>Se gonfler</i>	Subjective	Intransitive	Autocausative	Medio-passive	Medio-passive	Anticausative	Anticausative	Autocausative
<i>S'agir</i>	Subjective	(no data)	(no data)	Passive	(no data)	(no data)	Autonomous (syntax)	Impersonal passive
<i>Se passer</i>	Passive	Phantom agent	Quasi-passive	Passive	Passive	Mediopassive	Autonomous (syntax) Autonomous (selection)	Autocausative Agentless passive Impersonal passive
<i>Se construire</i>	Passive	Phantom agent	Quasi-passive	Passive	Passive	Mediopassive	Reflexive Mediopassive	Agentless passive Impersonal passive
<i>Se pouvoir</i>	Subjective	(no data)	(no data)	Passive	(no data)	(no data)	Autonomous (syntax)	Impersonal passive

Annex 2: Analysis of prototypical uses of pronominal verbs

Se laver

Le chat [...] se lavait l'oreille.

	<i>Le chat</i>	<i>se lavait</i>	<i>l'oreille</i>
<i>Semantic role</i>	A	V	P
<i>Syntactic function</i>	S	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: <i>Le chat lavait son oreille</i> ► Reduction of a syntactic argument (V cannot accept possessive)		

Je courus à la salle de bains et me lavai [...].

	<i>Je</i>	<i>me lavai</i>
<i>Semantic role</i>	A	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: <i>*Je lavai moi-même</i> ► Reduction of a syntactic argument (V cannot accept DO)	

	► Reduction of a semantic argument (V cannot accept P)
--	--

Le sang se lave avec les larmes [...].

	<i>Le sang</i>	<i>se lave</i>	<i>avec les larmes</i>
<i>Semantic role</i>	P	V	A
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>Les larmes lavent le sang</i> ► Demotion of a syntactic argument (S is expressed as CIRC) ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO)		

Se laver is a good example of a commonly used pronominal verb that can take different meanings according to the sentence it is used in. We could say that there is not one pronominal construction with « se laver », but at least three ; for each one, changes occur within the diathetic mapping. One common feature, however, is that there is a reduction in the valency of the verb when compared to a neutral diathesis variant (either the impossibility of adding a DO or a possessive).

In Kulikov's (2011) typology of diathesis, these three constructions correspond to the reflexive benefactive diathesis, the canonical reflexive diathesis and the conversive passive diathesis. The conversive passive, especially, is a less common diathesis where the original subject, instead of being deleted as it is often the case with passive diathesis, is turned into an oblique or a circumstantial. The patient still goes through the process of becoming the new subject, hence the belonging to the passive category.

Se regarder

Tout à l'heure, je me regardais dans la glace, [...].

	<i>Je</i>	<i>me regardais</i>	<i>dans la glace</i>
<i>Semantic role</i>	A	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>*Je regardais moi-même [...]</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)		

Même la mort des autres ne se regarde pas en face.

	<i>La mort des autres</i>	<i>ne se regarde pas</i>	<i>en face</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>On ne regarde pas la mort des autres en face</i>		

	► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)
--	--

[...] *Ils se regardaient dans les yeux* [...]

	<i>Ils</i>	<i>se regardaient</i>	<i>dans les yeux</i>
<i>Semantic role</i>	A (A1 & A2)	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>A1 regardait A2 dans les yeux et A2 regardait A1 dans les yeux</i> [...] ► Fusion of semantic arguments (A are put into one) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)		

Se regarder, just like *se laver*, is frequently used in a reflexive meaning, but also in its reciprocal acceptation. According to Kulikov's (2011) typology, the uses we uncovered for this verb correspond to the canonical reflexive diathesis, the agentless passive diathesis and the canonical reciprocal diathesis. The valency of the predicate is also reduced in all examples compared to the neutral diathesis.

The canonical reflexive diathesis seems to have a common mapping in which DO/P pair cannot be accepted anymore, implying the clitic *se* is taking their role both syntactically and semantically. This mapping is similarly found in the canonical reciprocal diathesis. However, in this particular diathesis, there is also a fusion of semantic arguments as the agent in the neutral variant should always be plural for this construction to be possible.

The agentless passive diathesis's main property is that the verb cannot accept a semantical agent, hence its calling. It also shares two valency changes commonly found in passive-type diathesis: the recession of a DO and the promotion of the neutral variant's DO to subject.

Se raser

Il se rase la barbe.

	<i>Il</i>	<i>se rase</i>	<i>la barbe</i>
<i>Semantic role</i>	A	V	P
<i>Syntactic function</i>	S	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: <i>Il rase sa barbe</i> ► Reduction of a syntactic argument (V cannot accept possessive)		

A présent, il faut que je me rase [...].

	<i>Je</i>	<i>me rase</i>
<i>Semantic role</i>	A	V

<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: * <i>Je rase moi-même</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)	

Se raser works mainly with two diathetic mapping: a reflexive benefactive diathesis and canonical reflexive diathesis. In the canonical reflexive diathesis, it is often implied one is shaving his beard, but it is not always the case as *Je me rase* is also used for other body parts (e.g. legs or armpits). It is likely that the reflexive benefactive diathesis for pronominal verbs like *se raser* has been abbreviated to form the meaning obtained with the canonical reflexive diathesis.

An analysis of the mapping of common diathesis within the sub-classification of reflexive verbs (*se laver*, *se regarder*, *se raser*) shows exact same patterns. It should be highlighted that the canonical reflexive diathesis, in particular, is found in each of these verbs and is likely to be a prominent mapping of pronominal-verbal constructions.

Reciprocal verbs

S'aimer

[...] *Je voulais que tout le monde s'aime* [...]

	<i>Tout le monde</i>	<i>s'aime</i>
<i>Semantic role</i>	A (A+A _α)	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: <i>Tout le monde aime tout le monde</i> [...] ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)	

S'aimer is mainly used for its reciprocal meaning, although some reflexive and passive uses are sometimes reported. Similarly to what happened in *Ils se regardaient dans les yeux*, the agent has to be plural (even though morphologically, *tout le monde* is a singular, its meaning is plural) for the reciprocal action to be performed. Also, DO/P couple is reduced but since *Tout le monde* already incorporates all entities of the process, the semantic fusion of agents is already present in the neutral variant.

Se voir

Anne et Pierre se sont vus hier soir [...].

	<i>Anne et Pierre</i>	<i>se sont vus</i>	<i>hier soir</i>
<i>Semantic role</i>	A (A1 + A2)	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>Anne a vu Pierre et Pierre a vu Anne</i> [...]		

	► Fusion of semantic arguments (A are put into one) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)
--	---

Il s'est vu récompenser à la fin de l'année scolaire.

	<i>Il</i>	<i>s'est vu</i>	<i>récompenser</i>
<i>Semantic role</i>	P	V	V
<i>Syntactic function</i>	S	V	V
<i>Case syntax</i>	NOM	V	V
<i>Valency change</i>	Neutral diathesis: <i>On l'a récompensé [...]</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a semantic argument (V cannot accept A)		

Se voir has a common reciprocal use working in the same fashion as *se regarder* or *s'aimer*.

Se voir + *infinitive* is a somewhat rare construction, mainly used in formal contexts, that has been thoroughly studied by Gmir-Ezzine (2014). According to the researcher, although this diathesis is following a similar pattern to the passive diathesis, this construction would correspond to a circumstantial diathesis, whereas unlike the passive, arguments other than DO can be promoted as subjects as seen in the following sentences:

Indirect object: *Il s'est vu décerner la légion d'honneur.* (Neutral: *On lui a décerné la légion d'honneur*)
Genitive: *La ville se voit conquérir ses terres par l'ennemi.* (Neutral: *Les terres de la ville sont conquises par l'ennemi*)

Since the mapping is each time slightly different, we could conclude that there are several diathesis based on *se voir* + *infinitive*. However, even if these uses are tied to this construction in particular, they do all share the property of promoting a circumstantial complement to subject; if we refer to Kulikov's typology (2011), these structures would then correspond to a converse passive.

Se parler

Elle se parlait à elle-même.

	<i>Elle</i>	<i>se parlait</i>	<i>à elle-même</i>
<i>Semantic role</i>	A	V	B
<i>Syntactic function</i>	S	V	IO
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: <i>*Elle parlait à elle-même à elle-même.</i> ► Reduction of a syntactic argument (IO reduced to one)		

Le français se parle.

	<i>Le français</i>	<i>se parle</i>
--	--------------------	-----------------

<i>Semantic role</i>	P	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: <i>On parle le français.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)	

Se parler à soi-même is a commonly used expression roughly equivalent to the English *to talk to oneself*. We can observe that in English and French alike, the emphatic *soi-même/oneself* is expressed. Nevertheless, French's pronominal construction allows to double down on the emphasis by using both the reflexive form and the emphatic. Since the reflexive *se parler* already derives from then neutral diathesis *parler à soi-même*, this construction turns out to be a pleonasm. Albeit rare, similar pleonastic uses can be found with other verbs with intransitive uses such as *se répondre* and *se mentir* (*Je me réponds à moi-même/Il se ment à lui-même*). In any case, this construction belongs to a particular case of the canonical reflexive diathesis where an IO is suppressed rather than a DO, because the original verb works as an indirect transitive.

However, *parler* also has a transitive use with a different meaning: *to speak (a language)*. In this acception, a case can be made for an agentless passive diathesis as seen in the second example *Le français se parle*.

Passive verbs

Se boire

Tisane qui se boit chaude [...]

	<i>Tisane qui</i>	<i>se boit</i>	<i>chaude</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>On boit cette tisane chaude.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)		

Se boire is often used in a diathesis where the agent is non-referential and irrelevant („whoever“), and the sentence takes the meaning of habituality, called potential agentless passive diathesis (Kulikov, 2011) or *SE-moyen* diathesis (Nakamura, 2014). In regard to the valency changes, since we are confronted to the exact same processes as with a regular agentless passive, this diathesis is merely considered as an extension of it due to semantic reasons.

Nevertheless, the pronominal variant allows the dummy, non-referential agent to be completely shut down, leading this construction to be commonly used in French.

Se vendre

Les chevaux [...] se vendaient vingt francs pièce [...]

	<i>Les chevaux</i>	<i>se vendaient</i>	<i>vingt francs pièce</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>On vendait les chevaux vingt francs pièce.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)		

Drumont [...] se vendait à cent milles, et son oeuvre fait le bruit d'une révolution.

	<i>Drumont</i>	<i>se vendait</i>	<i>à cent mille</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>On vendait Drumont à cent mille.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)		

Je me suis vendu corps et âme [...]

	<i>Je</i>	<i>me suis vendu</i>	<i>corps et âme</i>
<i>Semantic role</i>	A	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>*J'ai vendu moi-même corps et âme</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept P)		

These three examples are interesting in that they print each time a slight meaning derivation to the verb *se vendre*. The first two sentences are another examples of potential agentless passive diathesis, where the agent has no consistency whatsoever. However, while *vendre* literally means *to sell* in the first sentence, the meaning is extended as *to sell the work of* in the second phrase. Finally, the last sentence, which corresponds to a canonical reflexive diathesis, extends the meaning of *vendre* to *to sell one's qualities*. Even though the diathetic mappings are shared with the potential agentless passive and canonical reflexive diathesis we've seen up until now, the nuances taken by the verb should be considered carefully, as a given meaning would only be possible for a given diathesis, rather than for a given verb.

S'utiliser

L'intelligence artificielle s'utilise fréquemment [...]

	<i>L'IA</i>	<i>s'utilise</i>	<i>fréquemment</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: <i>On utilise fréquemment l'IA.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)		

S'utiliser is also a good example of verb often used for its nuance of habituality, consequently in potential agentless passive diathesis. The diathetic mapping corresponds to the standard already seen before with *se boire* and *se vendre*, which hints the coherence of this diathesis using pronominal verbs.

Subjective verbs

S'évanouir

La poussière [...] s'évanouit.

	<i>La poussière</i>	<i>s'évanouit</i>
<i>Semantic role</i>	E	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: <i>*La poussière évanouit elle-même</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)	

This construction, if we consider as plausible the rare transitive uses of *évanouir* (to make disappear, probably constructed from the pronominal variant), would actually work in a similar fashion as a canonical reflexive; however, this diathesis has its own characteristics, since the subject's semantic role is not an agent, but rather an experiencer (experiences the process of *s'évanouir* by itself, unlike the agent which conducts the process or the patient which undergoes the process conducted by an agent).

Labelle (2010) describes this diathesis as decausative because this is as a construction where there is no causal effect from an agent and the process is thus autonomous. Kulikov's (2011) typology refers to this diathesis as anticausative, which is coherent with the idea of removing any causal information linked to the verbal process; the anticausative „is consistent with the situation coming about spontaneously“ (Comrie, 1985). The anticausative/decausative diathesis is widely used for intransitive verbs designating a change of state, in which the subject plays the role of an experiencer (such as *se briser*, *se fossiliser*).

Se repentir

[...] Je me repents de ce que j'avais fait.

	<i>Je</i>	<i>me repents</i>	<i>de ce que j'avais fait</i>
<i>Semantic role</i>	E	V	C

<i>Syntactic function</i>	S	V	OBL
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: * <i>Je repents moi-même de ce que j'avais fait</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)		

Unlike with *s'évanouir*, there seems to be no historical trace of *repentir* used as a transitive verb (François, 2011). Semantically speaking, *se repentir* is roughly the equivalent of the non-pronominal verb *regretter* with possible intransitive uses (*Je regrette*). This verb also induces a change of emotional state, implying the subject is affected by the process while being at its origin, leading us to conclude that it takes the role of an experiencer, just like with *s'évanouir*. However, since the process is caused by an entity (*Je* in this case), this construction belongs to the autocausive diathesis rather than the anti/decausive (in which the cause of the process affecting the experiencer is not relevant/known).

Se douter

Je me doutais de tout ce que vous me dites.

	<i>Je</i>	<i>me doutais</i>	<i>de tout ce que vous me dites</i>
<i>Semantic role</i>	E	V	C
<i>Syntactic function</i>	S	V	OBL
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: * <i>Je doutais moi-même de tout ce que vous me dites</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)		

Se douter is verb of emotion with a strong peculiarity: its transitive variant, *douter*, is often understood as *to doubt* while the pronominal form takes the meaning of *to suspect*. While these two meanings are apparently antithetic (*I doubt he will come* vs *I suspect he will come*), Burston (1979) claims that the important is that both variants have in common the semantic nuance of 'lack of certainty', and that the clitic *se*, by including the subject as an experiencer rather than an agent, further induces the speaker to be mentally involved in the process. The difference between the two becomes more clear when comparing phrases with similar meaning: *Je doute qu'il viendra* „*I doubt he will come*“ vs *Je me doute qu'il ne viendra pas* „*I suspect he won't come*“. The lack of certainty is present in both cases, but the pronominal variant is rather seen as the conclusion of an evaluation, an internal process of thought while *douter* implies an automatic, unthinking reaction (Burston, 1979).

From a diathetical point of view, *se douter* functions similarly to *se repentir*, meaning the process occurs internally (without any external cause) and is experienced by the subject, hence the role of experiencer it takes. This is also an example of autocausive diathesis.

Se jouer

La bataille se jouera ici.

	<i>La bataille</i>	<i>se jouera</i>	<i>ici</i>
<i>Semantic role</i>	P	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: * <i>On jouera la bataille ici</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)		

La fortune se joue des hommes.

	<i>La fortune</i>	<i>se joue</i>	<i>des hommes</i>
<i>Semantic role</i>	E	V	C
<i>Syntactic function</i>	S	V	OBL
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: * <i>La fortune joue elle-même des hommes</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)		

Although intransitive verbs that have a pronominal form are rather rare (François, 2011) and *jouer* (along with *moquer*) is categorized by Zribi-Hertz (1987) as a intransitive verb, *jouer* still possesses some direct transitive uses (*jouer un morceau*, *jouer une somme*, *jouer un rôle*). Consequently, passive or reflexive pronominal constructions including *se jouer* are sometimes used (*Ce morceau se joue parfois dans ce bar*). The first sentence covered is another example of agentless passive diathesis, in which the verb *se jouer* has the meaning of *to take place/to be played*.

The second sentence describes a much more complicated use of *se jouer de*, often referred to as a literary variant of *jouer* and holding a meaning close to *se moquer de* (*to make fun of*). The diathetic mapping is also difficult to construct, since there is no obvious link between this pronominal use of *se jouer* and a neutral variant. However, a thorough look at the general semantics implied by *jouer* allows us to notice it is always related to the idea of recreation, fun. Since the semantic cluster represents an emotional state, the subject in the sentence is both at the origin of and affected by the process, hence the role of experiencer it takes. Indeed, when considering the subject as the experiencer, the pronominal construction makes much more sense: *La fortune se joue/s'amuse/se moque (du destin) des hommes*. Following Kulikov's (2011) typology, this diathesis would correspond to an autocausative diathesis.

Se mourir

Les institutions se meurent.

	<i>Les institutions</i>	<i>se meurent</i>
--	-------------------------	-------------------

<i>Semantic role</i>	E	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: * <i>Les institutions meurent elles-mêmes</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)	

Se mourir is often used as an alternative of *mourir* emphasizing on the durative aspect of the process. With *se mourir*, death (literal or figurative) is seen as a slow, painful phase. According to Burston (1979), rather than an organic, direct and brutal process like *mourir*, *se mourir* implies the active involvement of the subject, through complicity or resistance. This explains why the sentence *Les institutions se meurent* is imbued with some degree of self-destruction (institutions are participating in their own „death“) whereas *Les institutions meurent* simply states a fact. We can conclude, since the subject is also an experiencer, that the diathesis is autocausive.

Se casser

Hugo s'est cassé la jambe.

	<i>Hugo</i>	<i>s'est cassé</i>	<i>la jambe</i>
<i>Semantic role</i>	A	V	P
<i>Syntactic function</i>	S	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: <i>Hugo a cassé sa jambe</i> ► Reduction of a syntactic argument (V cannot accept possessive)		

Le pot se cassa en tombant.

	<i>Le pot</i>	<i>se cassa</i>	<i>en tombant</i>
<i>Semantic role</i>	E	V	C
<i>Syntactic function</i>	S	V	CIRC
<i>Case syntax</i>	NOM	V	ABL
<i>Valency change</i>	Neutral diathesis: * <i>Le pot cassa lui-même en tombant</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)		

Even though *casser* already has an intransitive usage with similar meaning, the form *se casser* is commonly used and holds several pronominal acceptions. From a diathetic point of view, these two sentences correspond to the reflexive benefactive diathesis and the anticausative diathesis. The first case *Hugo s'est cassé la jambe*²⁰ works similarly with *Le chat se lave l'oreille* that we mentioned earlier; while

²⁰ It is interesting to note the existence of the familiar expression *se casser*, which means *to leave* and is a derivation of *se casser la jambe*, which took a new meaning of *to run away* in the 19th century.

the verb does seem to encompass a DO, it actually loses the possessive which should be mandatory for the subject to be linked with the DO in the neutral variant. This kind of construction, often called body-part reflexives (Koenig, 1999), implies that the subject's action does something beneficial or detrimental to a part of himself rather than to its whole entity, hence the wording „reflexive benefactive diathesis“.

The second sentence is another example of anticausative diathesis. While the subject is not in total control of the action of breaking, the process is still marked by the absence of external cause (the pot simply fell and broke) and the subject undergoes a change of state (since it broke). We could thus consider the subject to be an experiencer, although determining whether it is a patient or an experiencer is not as clear when comparing *Le pot se cassa en tombant* with a sentence such as *Cette machine se casse facilement*. Even if there isn't any mentioned cause to the process in both cases, nothing implies that the subject undergoes the change of state by itself in the second sentence. As such, comparing *On casse cette machine facilement* vs **On cassa le pot en tombant* is a good indicator that the presence of *en tombant* gives some degree of autonomy to the process of the first sentence, hence the temptation to consider it as an experiencer. Nonetheless, using *se casser* in a potential agentless diathetic mapping (with a nuance of habituality) should not pose any problem either.

Se gonfler

La mer [...] se gonflait.

	<i>La mer</i>	<i>se gonflait</i>
<i>Semantic role</i>	E	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: * <i>La mer gonflait elle-même</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)	

The last verb we will cover is *se gonfler*. Just like *se casser*, this verb has an intransitive version with almost an equivalent meaning (*gonfler*). While the difference between *La mer gonflait* and *La mer se gonflait* is subtle, using the pronominal form *se gonfler* instead of *gonfler* has the effect of enhancing the focus on the change of state, rather than being a pure assertion (Labelle, 2010). From a diathetic point of view, there is no doubt that the subject of the pronominal sentence is an experiencer, since the process of *gonfler* can only happen internally by nature, and is not caused by anything else than a change of state. We can then conclude that this diathesis is autocausative.

Additional analysis

S'agir

Il s'agit de bijoux.

	<i>Il</i>	<i>s'agit</i>	<i>de bijoux</i>
<i>Semantic role</i>	Ø	V	C

<i>Syntactic function</i>	∇	V	OBL
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: * <i>Il agit lui-même de bijoux</i> ► Reduction of a syntactic argument (V cannot accept S) ► Reduction of a semantic argument (V cannot accept A)		

As an impersonal subject, *Il* is considered to be a dummy subject, which means it neither takes the role of an agent nor has any proper syntactic function in the process. However, since it logically takes the place of subject in the sentence, it is impossible to recover any subject/agent, and consequently to convert the sentence into a neutral diathesis. Kulikov's (2011) typology describes this type of construction as an impersonal passive diathesis with dummy subject.

S'agir is, in particular, a verb exclusively used in this type of diathesis and which has strongly drifted, in terms of meaning, from its non-pronominal variant *agir* (to act, to have an effect), although it could be argued that *s'agir* works as a passive form of *agir*; as seen in this old form (around 1600) from the TLFi: *Puisqu'il ne s'agit qu'à façonner des jardins, leur peu de contenu pourra s'accommoder sans dépense excessive = Si l'on agit pour façonner des jardins...*

A semantic analysis of the oblique complement *bijoux* shows that it is especially complex to determine the role it takes. Since *bijoux* does not seem to be an entity affected by any kind of process, but rather an observable element, we chose to consider this object as a circumstant (C). Indeed, the hypothesis of a patient would only be relevant if we chose to consider the verb as an actual active process; however, *il s'agit* works as a stative process (as it describes a state of being, cf. Michaelis, 2011), denying the patient hypothesis.

Se passer

Il se passe quelque chose.

	<i>Il</i>	<i>se passe</i>	<i>quelque chose</i>
<i>Semantic role</i>	Ø	V	C
<i>Syntactic function</i>	∇	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: * <i>Il passe lui-même quelque chose</i> ► Reduction of a syntactic argument (V cannot accept S) ► Reduction of a semantic argument (V cannot accept A)		

Just like with the impersonal passive *Il s'agit*, *Il se passe* is a stative process that denies the role of agent and marginalizes the function of subject. This verb works in pair with a DO rather than with an oblique, but still enables the consideration of the object as a circumstant from a semantic point of view. This sentence is also an example of impersonal passive with dummy subject.

Passer has a close semantic link with *se passer*, since both variants have the nuance of 'going through', although the pronominalization has the effect of taking a more figurative meaning. *Il se passe quelque chose* could then be roughly translated as *Something is going on [through this place]*.

Un appareil de ce genre [...] se passe d'aérodromes.

	<i>Un appareil</i>	<i>se passe</i>	<i>d'aérodromes</i>
<i>Semantic role</i>	E	V	C
<i>Syntactic function</i>	S	V	OBL
<i>Case syntax</i>	NOM	V	DAT
<i>Valency change</i>	Neutral diathesis: * <i>Un appareil passe lui-même d'aérodromes</i> ► Reduction of a syntactic argument (V cannot accept DO) ► Demotion of a semantic argument (A is expressed as E) ► Reduction of a semantic argument (V cannot accept P)		

Another common use of *se passer* can be found in this type of sentence. In this case, the meaning of *se passer* is 'to go without', for which a link could be made with the lexeme 'going through' of *passer* (in this case, *un appareil* is able to **go through** difficulties without using *un aérodrome*).

In this case, we believe that *Se passer* is an anticausative process, where the subject is both at the origin and affected by the process. The intrinsic meaning of the process of a sentence such as *Je me passe de chocolat* is comparable to *Je me repens de mon action*, because in both cases, the process is stative rather than active. Since there is no doubt that the process is caused by the internal properties of the subject *un appareil de ce genre*, the sentence belongs to the anticausative diathesis rather than anticausative.

Deux mois se sont passés [...].

	<i>Deux mois</i>	<i>se sont passés</i>
<i>Semantic role</i>	P	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V
<i>Valency change</i>	Neutral diathesis: <i>On a passé deux mois.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)	

This sentence represents another common use of *se passer*, which has roughly the same meaning as the one found in the impersonal passive. Unlike the impersonal passive, though, the subject is not a dummy one, but the patient of the process. Since the agent cannot be expressed in the sentence, it belongs to the agentless passive diathesis. The duality with the impersonal passive shows how flexible a pronominal verb can be in terms of structure.

Se construire

La maison s'est construite [...].

	<i>La maison</i>	<i>s'est construite</i>
<i>Semantic role</i>	P	V
<i>Syntactic function</i>	S	V
<i>Case syntax</i>	NOM	V

<i>Valency change</i>	Neutral diathesis: <i>On a construit la maison.</i> ► Promotion of a syntactic argument (DO is expressed as S) ► Reduction of a syntactic argument (V cannot accept DO) ► Reduction of a semantic argument (V cannot accept A)
-----------------------	---

This sentence is also an agentless passive similar to the one we just saw with *se passer*. *Construire* is a regular transitive verb which can be easily turned into its neutral variant; however, it requires an animate agent, so that the canonical reflexive interpretation is impossible with this sentence.

Il se construit beaucoup de maisons [...].

	<i>Il</i>	<i>se construit</i>	<i>beaucoup de maisons</i>
<i>Semantic role</i>	Ø	V	P
<i>Syntactic function</i>	∇	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: <i>On construit beaucoup de maisons</i> ► Reduction of a syntactic argument (V cannot accept S) ► Reduction of a semantic argument (V cannot accept A)		

While this structure is technically the same as *Il se passe quelque chose*, we figured out that it is important to think of *beaucoup de maisons* as a patient, because *se construire* is not a stative process, unlike *s'agir* or *se passer* at the impersonal passive. However, this structure keeps the same property of deleting the position of subject/agent and to replace it by a dummy subject. As such, and even though the diathetic mapping has a slight difference, it is still another case of impersonal passive diathesis with dummy subject. The coexistence of this structure with the ones found within stative pronominal verbs led us to consider the replacement of the subject/agent position by a dummy subject as this diathesis' main property.

Se pouvoir

Il se peut qu'il neige.

	<i>Il</i>	<i>se peut</i>	<i>qu'il neige</i>
<i>Semantic role</i>	Ø	V	C
<i>Syntactic function</i>	∇	V	DO
<i>Case syntax</i>	NOM	V	ACC
<i>Valency change</i>	Neutral diathesis: <i>*Il peut soi-même qu'il neige</i> ► Reduction of a syntactic argument (V cannot accept S) ► Reduction of a semantic argument (V cannot accept A)		

We would like to address this structure from a comparative manner with its non-pronominal impersonal variant *Il peut...* on basis of Rivière's (2018) analysis. The researcher makes a comparison of these two sentences:

Il peut neiger.

Il se peut qu'il neige.

Even though these two sentences show a lack of certainty, the first sentence doesn't involve a **subjective** thought on the possibility that the process might happen. *Se pouvoir*, on the other hand, indicates that the speaker believes that this possibility exists, rather than claiming an observable fact. This property can be collected in the following examples:

On peut construire un immeuble ici. (observable possibility)

Il se peut qu'on construise un immeuble ici. (subjective possibility)

From a diathetic point of view, the subject/agent is demoted as a dummy subject, and it is not even possible to recover any action in this sentence. As such, it is similar to *Il s'agit de bijoux* in the way that this construction is only a description of a situation. This sentence also belongs to the impersonal passive diathesis.

Annex 3: Pre-test

下の文を読んで正しい形を選んでください。

Dmesure predicted level: A1

Je [appelle/m'appelle] François et je suis belge.

フランスワと申します。ベルギー人です。

Elle va tous les jours à la gare à vélo. La gare [trouve/se trouve] près de sa maison.

彼女は毎日自転車で駅に行きます。駅は家の近くにあります。

La voiture de mon voisin [a vendu/s'est vendue] pour 3.000.000 yens il y a deux jours.

お隣さんの車は二日前に 300 万円で売れました。

Mon chat [mange/se mange] parfois du poisson.

私の猫はたまに魚を食べます。

Patrick et Marie [ont embrassé/se sont embrassés] sur un banc.

パトリックとマリはベンチでキスをしました。

Ton nouvel ordinateur [fonctionne/se fonctionne] très bien.

あなたの新しいパソコンは完璧に動きます。

La pièce « Hamlet » [a terminé/s'est terminée] la soirée.

「ハムレット」という劇は夜の章を終えました。

La France [a donné/s'est donnée] des tableaux de Monet au Japon.

フランスは日本に「モネ」の絵をあげました。

Lorsque nous avons visité le Louvre, nous [avons perdu/nous sommes perdus].

「ルーブル美術館」に行ったとき、道に迷いました。

L'école de ma fille [a mis/s'est mise] à distribuer des cahiers gratuitement.

娘の学校は無料でノートを配り始めました。

下の文の中に、フランス語で正しく書かれた文の前に **Oui** を付け、

間違って書かれた文の前に **Non** を付けてください。

Dmesure predicted level: A1

Thierry s'énerve facilement quand il fait chaud. 「énerver : 苛立たせる」

ティエリーは暑くなるとすぐに苛立ってしまいます。

Je vais à l'école en me prenant le bus. 「prendre le bus : バスに乗る」

バスに乗って学校に行きます。

Ce vin se boit avec du poisson. 「boire : 飲む」

このワインは魚と飲むべきです。

Au Japon, les voitures conduisent à gauche. 「conduire : 運転する」

日本では左側で車が運転されます。

Le soir, je mange un fruit avant de dormir. 「manger : 食べる」

夜は寝る前にフルーツを食べます。

Fiona et Jean se sont rencontrés au lycée. 「rencontrer : 会う」

フィオナとジョンは高校で出会いました。

Mon ordinateur s'allume très rapidement. 「allumer : 点ける」

私のパソコンはとても早く点けられます。

Le sucre décompose dans l'estomac. 「décomposer : 分解させる」

砂糖は胃の中で分解します。

Je m'écoute un morceau de Chopin. 「écouter : 聞く」

ショパンの曲を聞いてます。

Carlos s'est mort la semaine dernière. 「mourir : 亡くなる」

カルロスは先週亡くなりました。

下の動詞について代名動詞かどうかを判断し、相応しい形に活用してください。

(複数正答あり) (こちらの動詞の意味は「se」の意味を含みません)

Dmesure predicted level: A2

(Se) ranger 片付ける	Ce jouet dans cette boîte. このおもちゃはこの箱に片付けるべきです。
(Se) ranger 片付ける	Ma mère le jouet dans cette boîte. お母さんはおもちゃを箱に入れて片付けます。
(S')écouter 聞く	Anne Frank dans un message vocal. アンはフランクの声を音声メールで聞いてます。
(S')écouter 聞く	Frank et Anne dans un message vocal. アンとフランクはお互いの声を音声メールで聞いてます。
(S')adapter 適用させる	Je facilement aux changements de température. 温度変化に簡単に適用します。
(S')adapter 適用させる	Ce bus la climatisation selon la température. このバスは温度によってエアコンを適用させます。
(Se) passer 過ごす	Il plusieurs jours à l'hôpital. 何日か病院で過ごしました。
(Se) passer	Son opération sans problème.

過ごす	彼女の手術は問題なく終わりました。
(S')arranger 合わせる	Elle son agenda pour venir demain. 彼女は明日来れるように日程を合わせました。
(S')arranger 合わせる	Elle pour venir plus tôt demain. 明日早く来れるように努力します。
(S')illustrer 示す	Mon frère au tournoi de tennis. 兄がテニス大会で目立ちました。
(S')illustrer 示す	Le prix des balles un problème du tennis. ボールの値段はテニスの難題を示しています。
(Se) donner du temps 時間を与える	Je pour bien finir ce devoir. この宿題を終えるために自分に時間を与えます。
(Se) donner du temps 時間を与える	Le professeur nous pour finir ce devoir. この宿題を終えるために先生が時間を与えてくれました。
(Se) prendre un café コーヒーを飲む	Carlos avant d'aller travailler. カルロスは仕事に行く前にコーヒーを飲みます。

こちらのテキストを読んで空欄を埋めてください。

適切な動詞を選び、正しい形を書いてください。

(こちらの動詞の意味は「se」の意味を含みません)

Dmesure predicted level: A2

「(se) revoir ~ もう一度見る」「(se) parler ~ 話す」「(se) rassembler ~ 集める」「(s') envoyer ~ 送る」「(s') appeler ~ 呼ぶ」「(s') organiser ~ 開催する」「(se) manger ~ 食べる」「(se) décorer ~ 飾る」「(se) commencer ~ 始まる」「(se) déplier ~ 開く」「(se) téléphoner ~ 電話をかける」「(se) mettre ~ 置く」

Hier, Pierre est venu me voir pour[1]..... de l'anniversaire de Sophie. Il aimerait lui faire une surprise. Nous avons beaucoup discuté, mais nous n'avons pas trouvé quoi faire. Nous avons donc décidé de[2]..... ce matin. Finalement, aujourd'hui, nous[3]..... d'accord. De mon côté, je vais[4].....une invitation aux amis de Sophie et préparer un gâteau aux amandes. Ce gâteau[5]..... avec un petit café. Pendant ce temps-là, Pierre va[6]..... la maison avec des banderoles qui[7]..... facilement. Il va également[8]..... aux amis de Sophie pour qu'ils[9]..... avant de venir. J'ai hâte que ça[10]..... !

昨日、ソフィーの誕生日の話をするためにピエールが訪ねてきました。サプライズをしようとしています。色々話したけれど、どうすればいいか決められませんでした。それで今朝また会

うことにしました。結局意見が一つにまとまりました。私はソフィーのお友達に招待状を送ってアーモンドケーキを作ることになりました。このケーキはコーヒーで食べられます。その間、ピーエルが家の飾りつけをします。そのために、簡単に開ける横断幕を使います。そして、ソフィーの友達が集まるように電話をかけます。早く始まるといいね！

下の文において、代名動詞を使う理由を説明してください。
(分からなかったら空欄でいいです)

Dmesure predicted level: A2

Je m'appelle Yuki.

ゆきと申します。

Le verre s'est cassé malencontreusement.

グラスが何気に割れた。

Le chien de ma voisine s'est fait vacciner contre la rage.

お隣さんの犬に狂犬病のワクチンを打ってもらった。

Paul et Kevin se sont rencontrés au cinéma l'autre jour.

この間ポールとケビンが映画館で会った。

Fabio s'est pris un verre de vin pour accompagner son repas.

ファビオは食事のついでにワインを一杯取った。

Annex 4: Post-test

下の文を読んで正しい形を選んでください。

Dmesure predicted level: A1

Je 「rentre/me rentre」 à la maison tous les soirs vers 18h00.

毎日午後6時に家に帰ります。

La voiture de mon ami 「conduit/se conduit」 très facilement.

私の友達の車は簡単に運転出来ます。

On dit que les Espagnols 「couchent/se couchent」 souvent très tard.

スペイン人はよく遅く寝ると言われます。

François 「habite/s'habite」 près de la gare.

フランソアは駅の近くに住んでいます。

Marc 「a rencontré/s'est rencontré」 Alice dans un restaurant.

マルコはレストランでアリスと出会いました。

Mes nouveaux écouteurs 「achètent/s'achètent」 dans ce magasin.

私の新しいイヤホンはこの店で買えます。

La montgolfière 「a dégonflé/s'est dégonflée」 tout d'un coup avant de tomber.

気球が急にしぼんで落ちました。

Jean 「a préparé/s'est préparé」 le repas pour toute la famille.

ジョンは家族全員のためにご飯を用意しました。

Le lycée de mon frère 「trouve/se trouve」 près du parc Ohori.

弟の高校は大濠公園の近くにあります。
 Elle 「a perdu/s'est perdu」 la clé de son vélo.
 自転車の鍵を無くしました。

下の文の中に、フランス語で正しく書かれた文の前に **Oui** を付け、
 間違っている文の前に **Non** を付けてください。

Dmesure predicted level: A2

Quentin s'est choisi de partir au Canada. 「choisir : 決める」
 カンタンはカナダに行くことを決めました。
 J'aimerais bien goûter les huîtres de la région. 「goûter : 食べてみる」
 地方の牡蠣を食べてみたいです。
 Une bataille se prépare sur ces terres. 「préparer : 用意する」
 この土地で戦いが用意されている。
 J'aimerais me tomber amoureux rapidement. 「tomber : 落ちる」
 早く恋に落ちたいな。
 Les étoiles aident les voyageurs à guider. 「guider : 案内する」
 旅人は星の光に案内される。
 Le restaurant ferme pour le Nouvel An. 「fermer : 閉める、閉まる」
 お正月にレストランが閉まっている。
 J'ai vu un petit chien très mignon près du temple. 「voir : 見る」
 お寺の近くでとても可愛いワンちゃんを見た。
 La clé de ma voiture trouve sur la table. 「trouver : 見つける」
 車の鍵はテーブルに置いてある。
 Mes amies se sont rassemblées dans un café. 「rassembler : 集める」
 友達はカフェに集まった。
 La France s'est abolie la peine de mort. 「abolir : 廃止する」
 フランスは死刑を廃止した。

下の動詞について代名動詞かどうかを判断し、相応しい形に活用してください。
 (複数正答あり) (こちらの動詞の意味は「se」の意味を含みません)

Dmesure predicted level: A2

(Se) regarder	Les Américains beaucoup la télévision.
見る	アメリカ人はよくテレビを見る。
(Se) regarder	Je tous les matins dans le miroir.
見る	毎朝鏡で自分を見ている。
(Se) construire	L'entreprise de nouveaux bureaux ici.
建てる	会社はここで新しいオフィスビルを建てた。
(Se) construire	De nouveaux bureaux pour l'entreprise ici.

建てる	会社のための新しいオフィスビルはここに建てられる。
(S')accorder 与える	Le gouvernement de nouveaux droits aux commerçants. 政府は商人に新しい権利を与えた。
(S')accorder 与える	Je une pause avant de retourner étudier. 勉強を再開する前に休憩をとる。
(Se) vendre 売る	Il des journaux dans ce magasin. 彼はこの店で新聞を売っている。
(Se) vendre 売る	Ces journaux très rapidement. この新聞はすぐに売れる。
(S')égarer 無くす	Elle sa montre. 彼女は時計を無くしてしまった。
(S')égarer 無くす	Mon ami dans la forêt quand il était enfant. 友達は子供の頃森の中で道に迷った。
(Se) coincer 詰める	Mon vélo dans un fossé rempli de boue. 自転車は泥だらけの穴につかかった。
(Se) consulter 見てもらう	Mon fils un médecin régulièrement. 息子は定期的にお医者さんに見てもらっている。
(Se) consulter 見てもらう	Les partis à propos de ce problème. この問題に対してそれぞれの政党が話し合った。
(Se) tromper 騙す	Cet homme sa famille en cachant sa maladie. 病気を隠したこの男は家族を騙した。
(Se) tromper 騙す	Je rarement dans mes calculs. 計算に間違えることはめったにない。

こちらのテキストを読んで空欄を埋めてください。

適切な動詞を選び、正しい形を書いてください。

(こちらの動詞の意味は「se」の意味を含みません)

Dmesure predicted level: A1

「(s') écrire ~ 書く」「(se) retrouver ~ 取り戻す、合流する」「(s') aimer ~ ○が好き」「(se) voyager ~ 旅行する」「(s') énerver ~ 怒らせる」「(s') endormir ~ 眠かせる」「(se) continuer ~ 続ける」「(s') amuser ~ 楽しませる」「(se) prendre ~ 取る、乗る」「(se) promener ~ 散歩させる」「(se) savoir ~ 知る、分かる」「(se) casser ~ 壊す」

Avec Marie, nous sommes allées en voyage à Casablanca la semaine dernière. Nous
 [1]..... à l'aéroport avec tous nos bagages vers 8 heures du matin. Il a des magasins partout
 dans l'aéroport, alors nous avons décidé de [2]..... en attendant l'embarquement.
 J'achète souvent une bouteille d'eau pour le voyage, mais Marie achète plutôt des parfums.
 Elle [3]..... que c'est un peu moins cher sans les taxes. Ensuite, nous
 [4]..... l'avion. Je [5]..... dans l'avion, car j'étais fatiguée. Une fois
 arrivées, nous avons visité un peu la ville. Il y a un grand parc à Casablanca avec beaucoup de
 palmiers. Marie adore [6]....., alors elle [7]..... bien. Mais tout d'un
 coup, son sac a accroché une branche, et un des parfums qui était dans le sac est tombé
 et [8]..... . Marie [9]..... fortement . Elle
 [10]..... à boudier jusqu'à ce qu'on rentre à l'hôtel !

先週はマリーと一緒にカザブランカに行きました。朝8時ごろに荷物を持って空港で合流しました。空港にたくさんの店があって、搭乗までに散歩することにしました。旅行の前によく水を買います。しかし、無税で安くなると知っているマリーは香水を買います。そして、飛行機に乗りました。機内で疲れていたもので眠り込んでしまいました。着陸した後で街の観光をしました。カザブランカには、ヤシの木がたくさん並んでいる大きな公園があります。マリーは旅行するのが大好きで、とても楽しんでいました。しかし急に、カバンが枝に引っかかって、中にあった香水の瓶が落ちて壊れてしまいました。マリーは凄く怒り出しました。ホテルに戻るまでにずっと拗ね続けました！

下の文において、代名動詞を使う理由を説明してください。

(分からなかったら空欄でいいです)

Dmesure predicted level: A2

Je me suis lavé ce matin.

今朝お風呂に入った。

Les voitures japonaises se vendent facilement à l'étranger.

海外で日本製の車は簡単に売れる。

L'école de mon père s'est fait démolir il y a quelques mois par des ouvriers.

お父さんの学校は何か月前に取り壊された。

Le coût de la vie s'accroît de plus en plus.

物価がますます上がっている。

François et Marc se sont disputés hier.

昨日フランソアとマルクが喧嘩した。

Annex 5: Preliminary Survey

Variables: [1] Gender, [2] Age, [3] University, [4] Attitude and perception towards French and pronominal verbs, [5] Attitude and perception towards games, [6] Game type preference

Used as: co-variables for acquisition results

Format: [1] Multiple choice (MC), [2] Number, [3] Nominal, [4] [5] [6] 7-point Likert scale

Length: 1 MC item, 1 Number item, 1 Categorical item, 15 Likert items

VARIABLE: Biodata

- [1] 性別 男・女・他 [Gender – Male/Female/Other]
- [2] 年齢歳 [Age -years old (number)]
- [3] 大学名 [Name of University -]
- [4] Attitude and perception towards French and pronominal verbs
- {Experience} フランス語能力に自信がある [I feel confident in my language ability in French]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Experience} 代名動詞の使用に関して自信がある [I feel confident in using pronominal verbs]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Experience} 代名動詞の理解に自信がある [I feel confident in understanding pronominal verbs]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Importance} 代名動詞の習得が重要だと思う [I think learning pronominal verbs is important]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Interest} フランス語が楽しいと思う [I enjoy learning French]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Importance} 言語習得に文法の理解が重要だと思う [Grammatical understanding is important for language learning]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Importance} 代名動詞の習得に興味がある [I am curious about learning pronominal verbs]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Interest} フランス語能力を極めたいと思う [I look forward to improving my French skills]
- そう思わない 1 2 3 4 5 6 7 そう思う
- [5] Attitude and perception towards games
- {Interest} モバイルゲームが楽しいと思う [I enjoy playing mobile games]
- そう思わない 1 2 3 4 5 6 7 そう思う
- {Experience} モバイルゲームをよく使う [I often play mobile games]

全く使わない	1	2	3	4	5	6	7	頻繁に使う
{Experience} <u>教育のためのゲームをよく使う</u>				[I often use educative games]				
全く使わない	1	2	3	4	5	6	7	頻繁に使う
{Importance} <u>言語習得にゲームが便利だと思う</u>				[Games are helpful for learning languages]				
そう思わない	1	2	3	4	5	6	7	そう思う
{Importance} <u>学校以外にも言語習得が出来ると思う</u>				[Learning languages can also be done outside of school]				
そう思わない	1	2	3	4	5	6	7	そう思う
[6] Game type preference								
{Interest} <u>ゲームを使うときはストーリーを重視する</u>				[I like story-based games]				
嫌い	1	2	3	4	5	6	7	好き
{Interest} <u>短くて何回も出来るゲームが好き</u>				[I like short but repeatable games]				
そう思わない	1	2	3	4	5	6	7	そう思う

(The pictures used for the vocabulary size test can be accessed on request to the author)

Annex 6: Feedback Survey

Variables: [1] Perceived usefulness, [2] Perceived ease of use

Used as: co-variables

Format: [1] [2] 7-point Likert scale

Length: 12 Likert items

[1] Perceived usefulness								
<u>ゲームをした後、フランス語が得意になった</u>				[I am more confident in my language skill in French after playing the game]				
全然	1	2	3	4	5	6	7	非常に
<u>ゲームをした後、代名動詞の理解が深まった</u>				[I feel more confident in using pronominal verbs after playing the game]				
全然	1	2	3	4	5	6	7	非常に
<u>ゲームが役に立った</u>				[I feel like playing the game was useful]				
そう思わない	1	2	3	4	5	6	7	そう思う
<u>代名動詞をより使いこなせるようになった</u>				[I learned how to use pronominal verbs better]				
そう思わない	1	2	3	4	5	6	7	そう思う
<u>代名動詞をより早く認識出来るようになった</u>				[I am able to understand pronominal verbs more quickly]				
そう思わない	1	2	3	4	5	6	7	そう思う

代名動詞を使う必要がある状況を分かるようになった [I can recognise situations when a pronominal verb should be used]

そう思わない 1 2 3 4 5 6 7 そう思う

[2] Perceived ease of use

ゲームは使いやすかった [The game was easy to use]

そう思わない 1 2 3 4 5 6 7 そう思う

ゲームの内容をほとんど理解できた [I understood most of the game content]

そう思わない 1 2 3 4 5 6 7 そう思う

ゲームインターフェイスで困らなかった [I did not have trouble with the interface]

そう思わない 1 2 3 4 5 6 7 そう思う

ゲームで使われる言葉は分かりやすかった [The task wording was easy to understand]

そう思わない 1 2 3 4 5 6 7 そう思う

ゲームは自分が好きなように使えた [The game was flexible to use]

そう思わない 1 2 3 4 5 6 7 そう思う

ゲームに慣れるのが簡単だった [Becoming accustomed to the game was easy]

そう思わない 1 2 3 4 5 6 7 そう思う

Annex 7: Diathesis annotation guide (in French)

Guide d'annotation des structures pronominales en corpus

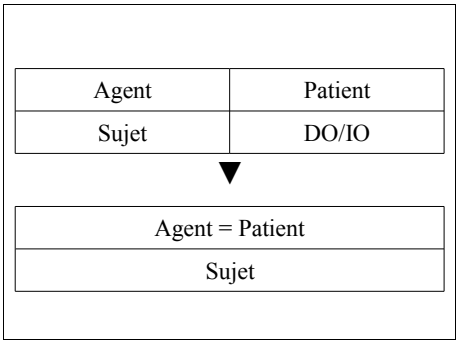
Présentation des diathèses représentées dans le corpus

Diathèses dynamiques

Diathèses conservant la valence de l'agent

Réflexif canonique

Les réflexifs canoniques suivent le schéma diathétique suivant:



Le patient de la variante transitive est remplacé par le clitique *se*. Ce patient peut être soit complément d'objet direct, soit complément d'objet indirect:

(1) *Ce sont déjà plus de 30 personnes qui se se sont inscrites.*

(2) *Je me parle parfois quand je suis concentré.*

On parle dès lors de réflexifs directs et de réflexifs indirects.

Ce schéma diathétique minimum peut

être suivi de divers compléments:

-Attribut du sujet:

(3) *Les hommes de Franck Renard, bien en jambes, se montrèrent plus sûrs.*

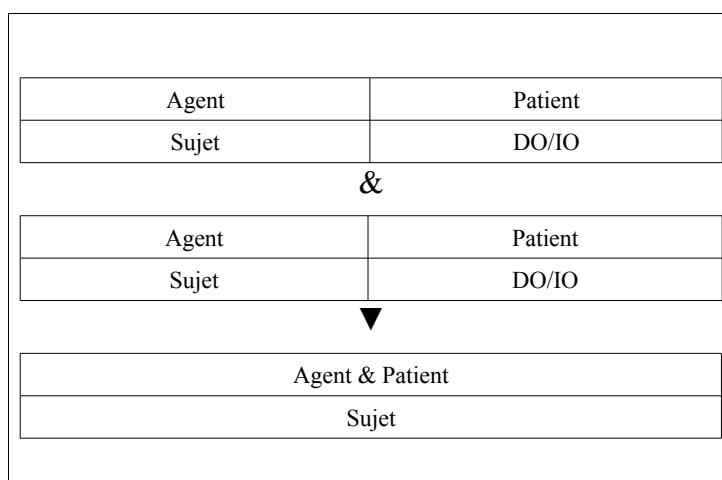
-Complément oblique:

(4) *Quelques enfants s'adonnaient à des glissades.*

(5) *Elle décide de rester à la maison pour s'occuper²¹ de sa fille.*

Réciproque canonique

Les réciproques canoniques suivent le schéma diathétique suivant:



Cette diathèse est caractérisée par la nécessité d'avoir, au moins **sémaniquement**, des agents multiples. On assiste à une fusion des rôles sémantiques qui interagissent entre eux. Comme pour les réflexifs, le patient des variantes transitives peut être soit complément d'objet direct, soit complément d'objet indirect:

(6) *La commission s'est réunie récemment.*

(7) *Ils se parlent régulièrement.*

Cette fusion des rôles sémantiques peut soit être implicite (verbes naturellement réciproques), soit provoquée par la présence du clitique *se*:

(8) *Ils se sont réunis au premier étage.*

(9) *Ils se sont regardés intensément.*

Creissels (2007) marque ici une comparaison entre les coopératifs (naturellement réciproques) et les réciproques à proprement parler. Cette distinction ne nous a pas semblé nécessaire étant donné que le schéma diathétique ne subit aucun changement et que la frontière entre coopératifs et réciproques est parfois compliquée à déterminer (*Ils se sont disputés* peut s'interpréter comme coopératif tout comme réciproque puisqu'on accepte également une variante au singulier *Il a disputé son amie*).

²¹ Occuper prend ici le sens de "employer" (TLFi)

Réflexif bénéfactif (ou autobénéfactif)

Les réflexifs bénéfactifs suivent le schéma diathétique suivant:

Agent		Patient
Sujet		DO
▼		
Agent	Patient	Bénéficiaire
Sujet	DO	IO
▼		
Agent & Bénéficiaire		Patient
Sujet		DO

Cette diathèse est caractérisée par l'ajout à la variante transitive d'un datif jouant le rôle de bénéficiaire. Ce complément est alors fusionné avec le sujet/agent, tout en maintenant un patient extérieur. Le processus décrit par un réflexif bénéfactif peut être soit positif et volontaire, soit négatif et involontaire:

(10) *Le conseil se réserve les pouvoirs.*

(11) *[...] les patients s'étant récemment fracturé la hanche.*

Notons que cette diathèse prendra priorité sur la réciprocité en cas de diathèse mixte (*Nous nous sommes donné rendez-vous* met en avant le rôle du datif de manière...réciproque).

Diathèses réduisant la valence de l'agent

Passif potentiel (sans agent)

Les passifs potentiels sans agent suivent le schéma diathétique suivant:

Agent	Patient
Sujet	DO
▼	
Patient	(Agent)
Sujet	-

Le passif potentiel sans agent réduit le rôle du patient de l'action qui devient sous-entendu, voire non-nécessaire. Le DO/patient de la variante transitive est alors promu sujet de la phrase:

(12) *La prudence s'impose chez les personnes âgées.*

(13) *Les projets se concrétisent petit à petit.*

L'utilité de cette diathèse réside dans les cas où l'importance du sujet/agent est réduit ("on", "n'importe qui"). C'est

notamment le cas dans des assertions modales (habitude ou devoir), où le rôle d'agent est naturellement moins grand que dans des actions ponctuelles:

(14) *Ce vin se boit frais.*

Cette diathèse peut exceptionnellement s'appliquer à des actions ponctuelles lorsqu'on désire mettre en avant le rôle passif de l'agent. Dans ce cas, on utilisera la formule causative *faire* + *infinitif*:

(15) *Jean s'est fait transporter dans sa chambre.*

Le fait que cette construction ne puisse s'appliquer qu'à des verbes agentifs (le conversif "*par* + complément d'agent" étant possible) constitue une preuve supplémentaire qu'on a bien à faire à un passif sans agent (Heidinger & Schäfer, 2008).

Passif conversif

Les passifs conversifs suivent le schéma diathétique suivant:

Agent	Patient
Sujet	DO
▼	
Patient	Agent
Sujet	OBL/CIRC

Tout comme le passif potentiel sans agent, le passif conversif met en avant le rôle du patient qui devient l'argument principal du verbe. Cette diathèse rend toutefois possible, dans le cas où l'agent initial a un lien sémantique fort avec le patient, de conserver l'agent dans la phrase sous forme de complément indirect ou oblique (Kulikov, 2011):

(16) *L'avion s'est rempli avec les passagers.*

(17) *La finale se jouera entre Utrecht et*

Soleuvre.

Cette diathèse, peu commune, permet de mettre en avant le rôle du patient/objet tout en gardant l'agent visible.

Le passif conversif n'est cependant pas toujours applicable:

(18) **La voiture s'est garée par/avec Jean*

(19) **Le chapeau se porte par/avec Jean*

La raison avancée par Kulikov (2011) est que la distance sémantique entre l'agent *Jean* et le patient *voiture* n'est pas assez réduite pour permettre une conversion de l'agent. Il n'y a effectivement pas de lien direct entre ces entités, là où *l'avion* et *les passagers* forment un couple évident, tout comme *la finale* et *(les équipes d') Utrecht et Soleuvre*.

Passif impersonnel

Les passifs impersonnels suivent le schéma diathétique suivant:

Agent	Patient
Sujet	DO
▼	
-	Patient
Factice	DO/OBL

Cette diathèse est marquée par la suppression de la fonction sujet/agent. Or, puisque le français est une langue qui n'autorise pas la suppression pure et simple de cette fonction, celle-ci est remplacée par un pronom impersonnel factice *il* (ou *ça/cela*). Ce pronom n'ayant aucun rôle sémantique, il sert uniquement à servir de point d'entrée syntaxique de la phrase. L'objet de la transitive devient alors argument

principal du verbe:

(20) *Il se construit beaucoup de ponts.*

(21) *Il se passe des choses nouvelles.*

(22) *Il ne s'agit pas d'un phénomène météorologique.*

Bien que la structure interprétative diffère quelque peu selon le verbe, on trouve toujours le même schéma diathétique avec un remplacement de la fonction sujet/agent par un pronom factice, suivi d'une construction objet (direct ou oblique).

Diathèses résultatives

Les diathèses suivantes sont formées à partir de verbes de nature résultative, ce qui signifie qu'ils mettent l'accent sur le résultat plutôt que sur le processus, à l'inverse des verbes dynamiques. Concrètement, on trouve dans cette catégorie les procès verbaux **non-agentifs**, c'est-à-dire qui décrivent des changements d'état exprimés sans cause (*se casser*, *se gonfler*, *s'éteindre*), des événements aspectuels (*se commencer*, *se dérouler*, *se passer*), des caractérisations (*se définir*, *se manifester*), des états physiques (*se reposer*, *se prélasser*), des états émotionnels (*s'intéresser*, *se scandaliser*), des états cognitifs (*s'imaginer*, *s'apercevoir*, *se douter*). Les verbes non-agentifs ne se basent pas sur la dynamique entre un **Agent** + **Patient** mais plutôt entre une **Cause** + **Expérient/Thème** (Kulikov, 2011).

Autocausatif

Les autocausatifs suivent le schéma diathétique suivant:

Cause	Expérient	(Patient)
Sujet	DO	(OBL)
▼		
Cause ≈ Expérient		(Patient)
Sujet		(OBL)

Prototypiquement, le sujet d'une construction autocausative est à la fois considéré comme "cause" (stimulus) et "expérient" (lieu de manifestation) d'un procès. La différence entre un réflexif canonique et l'autocausatif réside dans la sémantique du verbe, qui ne permet pas de considérer le sujet de la variante transitive comme agent sous contrôle et conscient d'une action. Toutefois, l'autocausatif indique que le procès verbal reste le résultat de l'implication,

au moins partielle, d'un sujet actif (Creissels, 2007):

(23) *Je me suis aperçu de mon erreur.*

(24) *Je me repents de mes actions.*

Les verbes d'émotions, en particulier, sont régulièrement autocausatifs puisqu'ils représentent typiquement des événements internes mais pas complètement contrôlés et/ou conscients:

(25) *Ils se sont intéressés à l'expression orale.*

(26) *L'homme s'est frustré de son impuissance.*

Puisque le sujet de l'autocausatif est en partie acteur du procès, il sera en général animé.

Anticausatif

Les anticausatifs suivent le schéma diathétique suivant:

Cause	Patient/Thème
Sujet	DO
▼	
Patient/Thème	(Cause)
Sujet	-

Contrairement aux autocausatifs, le procès verbal représentatif des anticausatifs est entièrement involontaire, incontrôlé et, dans ce cas, forcément lié à des causes extérieures. Le contexte de la phrase ne permettant pas de déterminer la cause, celle-ci est simplement omise et le sujet est affecté par un procès qui semble spontané. Il est impossible pour un vase de se briser tout seul, même si cette impression en ressort dans (27):

(27) *Le vase s'est brisé tout d'un coup.*

(28) *L'os se cassa.*

(29) *Un accident s'est produit.*

Comme nous pouvons le voir, nous pouvons tout à fait trouver des verbes agentifs dans cette catégorie lorsqu'ils sont dénués de causalité (*se briser*, *se casser*). Cela fait qu'on utilise parfois l'appellation "Thème" au lieu de Patient pour différencier le sujet des anticausatifs du sujet des passifs potentiels sans agent.

A l'inverse de l'autocausatif, l'anticausatif s'appliquera typiquement à des êtres inanimés, ce qui fait du sujet un thème fonctionnant sans agent impliqué:

(30) *La séance s'est terminée à minuit.*

(31) *La porte s'est ouverte toute seule.*

La différence entre anticausatifs et passifs sans agent est subtile et tient au fait que le passif sans agent est un renversement d'une tournure transitive active, où un agent, même générique, est considéré comme responsable de l'action sur le patient, là où l'anticausatif condamne a priori l'intervention d'un acteur sur le procès qui paraît spontané, évident. C'est la raison pour laquelle le procès d'un anticausatif ne peut pas être agentif.

Méthodologie d'annotation du corpus

Les diathèses majeures s'appliquant aux pronominaux ayant été présentées dans la section précédente, il convient à présent de décrire la procédure à employer pour attribuer une diathèse à une structure donnée.

Première étape: Identifier le sens du verbe

La première étape consiste à identifier le sens pris par le procès verbal dans la construction qu'on tente d'analyser. Le français possédant de nombreux verbes hautement polysémiques, un travail de recherche étymologique et/ou de linguistique comparative peut être nécessaire pour appréhender certaines formes qui semblent autrement difficilement analysables:

(32) *Elle décide de rester à la maison pour s'occuper de sa fille.*

Dans cet exemple, on pourrait d'abord imaginer que *sa fille* a le rôle de patient du verbe *s'occuper*. Toutefois, une analyse étymologique dans le TLFi ainsi qu'une comparaison avec l'italien *occuparsi* permet de mieux comprendre la pronominalisation. On trouve ainsi la forme *occupar quelqu'un de quelque chose*²², qui correspond au sens transitif pris par *s'occuper* dans (30). La fréquence d'utilisation d'un verbe n'est pas prise en compte pour cette étape, pour autant qu'il soit attesté historiquement, dans le langage courant ou sur le web. En effet, il est impossible de vérifier si une utilisation historiquement attestée est réellement dénuée d'usages contemporains.

Deuxième étape: Déterminer le schéma diathétique

La deuxième étape demande d'identifier les principaux arguments de la structure pronominale en terme de fonctions syntaxiques et de rôles sémantiques. Ce travail interprétatif est essentiel pour déterminer à quelle diathèse correspondra la structure en question. Un inventaire des principales fonctions syntaxiques et des rôles sémantiques (cf. Carnie, 2007; Thomas, 2007) a été utilisé:

Fonction syntaxique	Explication	Exemple
Sujet (S)	Objet régissant le verbe dans la phrase	<i>Jean mange la pomme.</i>
Objet direct (DO)	Objet lié directement à un verbe transitif	<i>Jean mange la pomme.</i>
Objet indirect (IO)	Objet lié par une préposition à un verbe transitif	<i>Jean offre un cadeau à Marie.</i>
Objet oblique (OBL)	Objet lié par une préposition à un verbe intransitif	<i>Jean joue au poker.</i>
Datif (DAT)	Objet destinataire du verbe	<i>Marc écrit pour sa mère.</i>
Génitif (GEN)	Objet attributif d'un autre objet	<i>Les oreilles du chat sont petites.</i>
Circonstanciel / Adjoint (CIRC/ADJ)	Objet externe (non-argument) descriptif	<i>Jean mange la pomme sur un banc.</i>
Factice (∇)	Fonction factice (cf. les impersonnels <i>Il, ça</i>)	<i>Il pleut.</i>

22 *Ces hésitations l'occupaient du docteur* (ZOLA, Page amour, 1878, p.811)

Rôle sémantique	Explication	Exemple
Agent (A)	Entité réalisant volontairement le procès verbal	<i><u>Jean</u> mange la pomme.</i>
Patient (P)	Entité subissant le procès verbal	<i>Jean mange <u>la pomme</u>.</i>
Bénéficiaire (B)	Entité bénéficiant du procès (positivement ou négativement)	<i>Jean offre un cadeau à <u>Marie</u>.</i>
Expérient (E)	Entité animée affectée par un procès verbal sensoriel ou émotionnel	<i><u>Jean</u> a peur.</i>
Thème (T)	Entité inanimée affectée par un procès verbal sans cause identifiable	<i><u>La pluie</u> a repris de plus belle.</i>
Instigateur (I) / Causant (CAUS)	Entité animée à l'origine du procès verbal	<i><u>Marc</u> a fait venir Jean.</i>
Cause (CAUS) / Stimulus (STI)	Entité inanimée à l'origine du procès verbal	<i><u>La pluie</u> a fait annuler la représentation.</i>
Circonstanciel / Adjoint (C/ADJ) (locatif, manière, etc.)	Entité détaillant les circonstances du procès verbal	<i>Jean mange la pomme <u>sur un banc</u>.</i>
Factice (Ø)	N'a aucun référent sémantique	<i><u>Il</u> pleut.</i>

L'analyse du sens du verbe lors de la première étape offre une aide à cette étape d'identification. Dans le cas de (32), on va donc identifier les fonctions syntaxiques/rôles sémantiques suivants:

Elle = Sujet/Agent

s' = Sujet/Patient

occuper = Verbe (transitif direct)

sa fille = OBL/Bénéficiaire

Troisième étape: protocole de décision

Le protocole de décision se fait en répondant à des questions binaires, qui servent de "tests" à la manière d'un organigramme de programmation. L'ordre de ces questions est important, puisqu'elles partiront des distinctions les plus essentielles aux plus précises. Les questions suivantes sont ainsi posées:

1) Le sujet est-il un pronom impersonnel factice?

Oui ► **passif impersonnel**

Non ► 2)

2) Le procès verbal est-il dynamique ou résultatif?

Dynamique ► 3a)

Résultatif ► 3b)

Un procès verbal dynamique met en évidence une action où un agent (A) agit en contrôle et de manière volontaire sur un patient (P).

Un procès verbal résultatif met en évidence un état ou un changement d'état affectant le sujet; il implique qu'il n'y a généralement pas d'action réalisée entre un agent (A) et un patient (P).

Si le procès verbal est dynamique

3a) Le sujet syntaxique est-il également agent (volontairement à l'origine du procès)?

Oui ► 4a)

Non ► 4b)

4a) Le verbe sous forme pronominale a-t-il un objet (direct/oblique) et le clitique est-il un datif?

Oui ► **bénéfactif**

Non ► 5a)

5a) Le procès verbal correspond-il à une interaction entre plusieurs agents (directement ou pas)?

Oui ► **réciproque**

Non ► **réflexif**

4b) L'agent est-il exprimé sous forme d'objet oblique/circonstantiel?

Oui ► **passif conversif**

Non ► 4c)

4c) L'action apparaît-elle comme spontanée et due à des causes inconnues?

Oui ► **anticausatif**

Non ► **passif sans agent**

Si le procès verbal est résultatif

3b) Le procès verbal porte-t-il sur une autre entité que le sujet?

Oui ► **bénéfactif**

Non ► 4d)

4d) Le procès verbal est-il, en partie ou totalement, provoqué par le sujet lui-même?

Oui ► **autocausatif**

Non ► **anticausatif**

Quatrième étape (optionnelle): vérifier via des tests

En cas de doute ou pour vérifier la pertinence de son choix, on peut effectuer des tests spécifiques à chaque construction diathétique. Ces tests sont principalement retenus des travaux de Creissels (2007).

Test du réflexif canonique:

Paraphrase **A V que moi-même/A V qu'à moi-même**

Mettre au pluriel + paraphrase **mutuellement**

Je me défends = Je ne défends que moi-même/Ils se défendent mutuellement

Je me parle = Je ne parle qu'à moi-même/Ils se parlent mutuellement

Je m'accroche à la bouée = Je n'accroche que moi-même/Ils s'accrochent mutuellement

Test du réciproque canonique:

Paraphrase **A V B & B V A** → réciproque simple

Paraphrase **A se V avec B & B se V avec A** → réciproque coopératif

Nous nous aimons = Je t'aime & Tu m'aimes

Nous nous entendons bien = Je m'entends bien avec toi & Tu t'entends bien avec moi

Nous nous battons = Je me bats avec toi & Tu te bats avec moi

Test du réflexif bénéfactif:

Paraphrases **A V B/A V possessif B** (l'une ou l'autre, les deux)

Je m'achète un cadeau = J'achète un cadeau/J'achète mon cadeau

Je me prends une amende = Je prends une amende

Je me suis cassé le bras = J'ai cassé mon bras

Test du passif sans agent:

Paraphrases **on doit Vinf A/on peut Vinf A** (l'une ou l'autre, les deux)

Ce vin se boit avec du fromage = On doit/peut boire ce vin avec du fromage

Mon appartement se trouve ici = On peut trouver mon appartement ici

Le français se parle en Belgique = On doit/peut parler français en Belgique

Test du passif conversif:

Paraphrase **B V A**

Ce livre se vend chez les connaisseurs = Les connaisseurs vendent ce livre

Ce traitement s'est testé avec ces patients = Ces patients ont testé ce traitement

Test du passif impersonnel:

Constructions spéciales du type **Il s'agit de, Il se peut que, Il s'avère que**

Paraphrase **B se V**

Il se construit des ponts ici = Des ponts se construisent ici

Il se passe des choses étranges = Des choses étranges se passent

Test de l'autocausatif:

Paraphrase **A V que moi-même/A V qu'à moi-même** → pas interprétable

Mettre au pluriel + paraphrase **mutuellement** → interprétation différente

Je me souviens (état cognitif) ~? ~~**Je ne souviens que moi-même**~~

Je m'ennuie (état émotionnel) ~ **Ils s'ennuient mutuellement** (action)

Je me rappelle (état cognitif) ~ **Ils se rappellent mutuellement** (action)

Je me repose (état physique) ~? ~~**Je ne repose que moi-même**~~

Test de l'anticausatif:

Paraphrases **on doit Vinf A/on peut Vinf A** → pas interprétable

La douleur se manifeste rapidement ~? ~~**On doit/peut manifester la douleur rapidement**~~

Mon audition s'est bien passée ~? ~~**On a dû/pu bien passer mon audition**~~

Cet os s'est cassé avec la pression ~? ~~**On a dû/pu casser cet os avec la pression**~~

La pluie s'est arrêtée abruptement ~? ~~**On a dû/pu arrêter la pluie**~~