The role of metalinguistic awareness and of L2 proficiency in positive lexical transfer from English (L2) to German (L3) by French-speaking Quebeckers

Thèse

Nina Woll

Doctorat en linguistique
Philosophiae doctor (Ph.D.)

Québec, Canada

The role of metalinguistic awareness and of L2 proficiency in positive lexical transfer from English (L2) to German (L3) by French-speaking Quebeckers

Thèse

Nina Woll

Sous la direction de :

Kirsten Hummel, directrice de recherche
Zita de Koninck, codirectrice de recherche

RESUME

Il a été avancé que des apprenants expérimentés développeraient des niveaux élevés de conscience métalinguistique (MLA), ce qui leur faciliterait l'apprentissage de langues subséquentes (p.ex., Singleton & Aronin, 2007). De plus, des chercheurs dans le domaine de l'acquisition des langues tierces insistent sur les influences positives qu'exercent les langues précédemment apprises sur l'apprentissage formel d'une langue étrangère (p.ex., Cenoz & Gorter, 2015), et proposent de délaisser le regard traditionnel qui mettait l'accent sur l'interférence à l'origine des erreurs des apprenants pour opter pour une vision plus large et positive de l'interaction entre les langues. Il a été démontré que la similarité typologique ainsi que la compétence dans la langue source influence tous les types de transfert (p.ex., Ringbom, 1987, 2007). Cependant, le défi méthodologique de déterminer, à la fois l'usage pertinent d'une langue cible en tant que résultat d'une influence translinguistique (p.ex., Falk & Bardel, 2010) et d'établir le rôle crucial de la MLA dans l'activation consciente de mots ou de constructions reliés à travers différentes langues, demeure.

La présente étude avait pour but de relever ce double défi en faisant appel à des protocoles oraux (TAPs) pour examiner le transfert positif de l'anglais (L2) vers l'allemand (L3) chez des Québécois francophones après cinq semaines d'enseignement formel de la L3. Les participants ont été soumis à une tâche de traduction développée aux fins de la présente étude. Les 42 items ont été sélectionnés sur la base de jugements de similarité et d'imagibilité ainsi que de fréquence des mots provenant d'une étude de cognats allemands-anglais (Friel & Kennison, 2001). Les participants devaient réfléchir à voix haute pendant qu'ils traduisaient des mots inconnus de l'allemand (L3) vers le français (L1). Le transfert positif a été opérationnalisé par des traductions correctes qui étaient basées sur un cognat anglais. La MLA a été mesurée par le biais du THAM (Test d'habiletés métalinguistiques) (Pinto & El Euch, 2015) ainsi que par l'analyse des TAPs. Les niveaux de compétence en anglais ont été établis sur la base du Michigan Test (Corrigan et al., 1979), tandis que les niveaux d'exposition ainsi que l'intérêt envers la langue et la culture allemandes ont été mesurés à l'aide d'un questionnaire.

Une analyse fine des TAPs a révélé de la variabilité inter- et intra-individuelle dans l'activation consciente du vocabulaire en L2, tout en permettant l'identification de niveaux

distincts de prise de conscience. Deux modèles indépendants de régressions logistiques ont permis d'identifier les deux dimensions de MLA comme prédicteurs de transfert positif. Le premier modèle, dans lequel le THAM était la mesure exclusive de MLA, a déterminé cette dimension réflexive comme principal prédicteur, suivie de la compétence en anglais, tandis qu'aucune des autres variables indépendantes pouvait prédire le transfert positif de l'anglais. Dans le second modèle, incluant le THAM ainsi que les TAPs comme mesures complémentaires de MLA, la dimension appliquée de MLA, telle que mesurée par les TAPs, était de loin le prédicteur principal, suivie de la dimension réflexive, telle que mesurée par le THAM, tandis que la compétence en anglais ne figurait plus parmi les facteurs ayant une influence significative sur la variable réponse. Bien que la verbalisation puisse avoir influencé la performance dans une certaine mesure, nos observations mettent en évidence la contribution précieuse de données introspectives comme complément aux résultats basés sur des caractéristiques purement linguistiques du transfert. Nos analyses soulignent la complexité des processus métalinguistiques et des stratégies individuelles, ce qui reflète une perspective dynamique du multilinguisme (p.ex., Jessner, 2008).

ABSTRACT

It has been posited that experienced language learners develop higher levels of metalinguistic awareness (MLA), which, in turn, appears to facilitate the acquisition of subsequent languages (e.g., Singleton & Aronin, 2007). Moreover, researchers in the field of Third Language Acquisition (TLA) emphasize the positive influences of previously acquired languages in foreign language classrooms (e.g., Cenoz & Gorter, 2015), in contrast to the traditional focus on interference-based learner errors in the study of language transfer. Typological similarity and source language proficiency are known to influence transfer processes of any kind (e.g., Ringbom, 1987, 2007). However, it remains a methodological challenge both to identify the felicitous use of a target language feature as an effect of cross-linguistic influence (e.g., Falk & Bardel, 2010) and to clearly determine the role of MLA for the conscious activation of related words or constructions across languages.

The present study aimed at meeting this double challenge by using think-aloud protocols (TAPs) to investigate positive lexical transfer from English (L2) to German (L3) by French-speaking Quebeckers (N = 66) after five weeks of formal L3 instruction. Participants completed a translation task consisting of 42 items selected on the basis of similarity and imageability ratings as well as word frequencies (Friel & Kennison, 2001). Participants were asked to think aloud while attempting to translate these largely unknown words from German (L3) into French (L1). Positive transfer was operationalized by correct translations that were related to an English cognate. Levels of MLA were measured by means of the THAM (Test d'habiletés métalinguistiques) (Pinto & El Euch, 2015) and complemented by the analysis of the TAPs. Proficiency levels in English were established by means of the Michigan Test (Corrigan et al., 1979), whereas exposure to the different languages as well as interest in the German language and culture were addressed in a background questionnaire.

A fine-grained analysis of the TAPs revealed inter- and intra-individual variability in the conscious activation of related L2 vocabulary while allowing for an identification of distinct levels of awareness. Two independent models of logistic regressions revealed that both dimensions of MLA significantly predicted positive transfer. The first model, which featured the THAM as the exclusive measure of MLA, identified this reflexive dimension

as the strongest predictor, followed by English language proficiency, whereas none of the other independent variables predicted positive transfer from English. In the second model, which included both the THAM and the TAPs as complementary measures of MLA, the applied dimension of MLA, as measured by the TAPs, was by far the strongest predictor, followed by the reflexive dimension, measured by the THAM, while English proficiency was no longer a significant predictor. Despite the fact that verbalization may have influenced overall performance to a certain extent, our observations point to the invaluable contribution of introspective data to complement findings based on language-inherent characteristics of crosslinguistic influence. Our analyses highlight the complexity of metalinguistic processes and individual strategies related to learner and context variables, in line with a dynamic view of multilingualism (e.g., Jessner, 2008a).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1 Introduction and research problem

With the growing mobility of the world population, the mastery of several languages is often regarded as a prerequisite and, above all, as an advantage in many areas of society. Consequently, a number of questions have been raised in the field of second language acquisition (SLA) research with respect to the multiplication of linguistic systems in the learner's mind. An increasing interest in the specific developments related to the acquisition of languages beyond the second has led to the emergence of a distinct research area. More specifically, research in third language acquisition (TLA) focuses on the complexity of the processing mechanisms that characterize learners of multiple languages (e.g., Aronin & Singleton, 2012; De Angelis, 2007; de Bot, 2004; García-Mayo, 2012; Herdina & Jessner, 2002). In more traditional perspectives on second and foreign language acquisition, in which research was primarily focused on contrasts between native and non-native languages (Lado, 1957; Weinreich, 1953), the focus was on an increased potential for interferences. Interestingly, however, rather than being disadvantaged with respect to comprehension and production processes in their additional languages, users of three or more languages seem to develop an increased flexibility with respect to the various challenges that language learning involves (e.g., Jessner, 2006). In line with Bialystok's findings on the cognitive development of bilinguals in the field of developmental psychology (e.g., Bialystok, 2001a, 2009), research in psycholinguistics has increasingly focused on the cognitive and affective factors that provide experienced language learners with enhanced abilities for further language learning, especially in the context of TLA, where both native and non-native background language(s) were shown to influence the acquisition process of a given L3[1] in different ways (e.g., Hufeisen & Gibson, 2003; Jessner, 1999; Singleton & Aronin, 2007).

---

[1] In line with Hammarberg's (2010) terminology, the term L3 will be used to denote all additional languages beyond the second, thus implying that TLA is concerned with experienced learners of at least one non-native language who are currently engaged in the process of learning an additional language, the L3 (even if, chronologically speaking, it could be considered their L4, etc.).

In Europe, authorities in educational politics have added multiple language acquisition to their agenda, usually with regard to intercultural competence (e.g., the Common European Framework of Reference for Languages, Council of Europe, 2001). Moreover, the so-called "intercomprehension" movement (European Commission, 2008; Hufeisen, 2000) has inspired new teaching methods, such as, for instance, the *EuroCom* project (Hufeisen & Marx, 2007), to promote the active use of background languages in the foreign language classroom. In North America, research in this area seems to be rather scarce, especially with regard to multiple language acquisition. Still, it has to be acknowledged, that in Canada, learning more than one non-native language has also become more widespread. In the province of Québec, for instance, most teenagers will have received ten years of English (L2) instruction when they graduate from high school. Some of them will also have started to learn Spanish as a third language at the secondary level. Those who choose to enter a junior college can then learn German as well as Spanish and in some institutions, Mandarin. It thus seems equally important to consider fostering *intercomprehension* in Canadian classrooms.

In the light of the above observations, it seems critically important to examine the potential benefits of crosslinguistic interaction for the acquisition of third or additional languages. The present study is situated in the field of TLA and addresses the interaction between two non-native languages, namely English and German, of French-speaking Quebeckers. When elements from one language influence the comprehension or production of another, researchers generally refer to "language transfer" (e.g., Odlin, 1989; Selinker, 1969) or "cross-linguistic influence" (e.g., Odlin, 2003). On the one hand, such processes can hinder comprehension (e.g., false cognates) or result in production errors or language switches, usually referred to as "interferences" or instances of "negative transfer" (Selinker, 1969, p. 91; see also Weinreich, 1953). On the other hand, instances in which another language provides cues that facilitate the comprehension or production of target language elements are referred to as "positive transfer" (Jarvis & Pavlenko, 2008, p. 25; Odlin, 2003, p. 438). Given the particular conditions of multiple language acquisition that were alluded to above, it is mostly with the emergence of TLA over the past fifteen years that positive crosslinguistic influences have come to be investigated more thoroughly, especially in

Europe (Gibson & Hufeisen, 2003; Marx & Mehlhorn, 2010; Odlin & Jarvis, 2004; Peyer, Kaiser, & Berthele, 2010; Rast, 2010). However, the interaction between different languages in the learner's mind can also lead to perturbation and research in applied linguistics must and still does account for interferences such as language switches in L2 or L3 production (e.g., Burton, 2013).

When it comes to identifying factors leading to transfer, either positive or negative, one of the most prominent seems to be the typological proximity (or distance) between the languages involved (e.g., Cenoz, 2003b; Odlin & Jarvis, 2004; Williams & Hammarberg, 1998). Moreover, the similarities and differences between languages that are merely perceived as such by the learner have been identified as equally important for transfer phenomena as real typological cues. This phenomenon was identified by Kellerman (1983) as "psychotypology" and examined by a number of researchers in TLA (e.g., De Angelis & Selinker, 2001; Ringbom, 2007). Other factors that are not related to purely linguistic aspects of the languages involved but rather to specific learner profiles are levels of language proficiency and particular features of language use. As far as language proficiency is concerned, it is not surprising that transfer is more likely to occur in low-proficiency languages, given that gaps in the target system tend to be overcome by support from either the L1 or, in case of TLA, from a given L2. At the same time, the proficiency level in the respective source language also seems to be a predictor for crosslinguistic influence. In other words, learners with higher levels of proficiency in a given L2 are more likely to draw upon this linguistic knowledge while learning a given L3 (e.g., Hammarberg, 2001; Odlin, 2003). Moreover, the frequency and recency of use of the L2 may further influence the degree to which the latter will be available to a learner when performing linguistic tasks in an L3 (e.g., Dewaele, 1998, 2001; Williams & Hammarberg, 1998). In line with these findings, it has been further posited that L2 exposure may enhance L2 activation in L3 production to an even higher extent than L2 proficiency (Tremblay, 2006).

Most of the above-mentioned studies have focused on non target-like language production, whereas relatively little has been said about the ways in which L2 knowledge might in fact help L3 learners. The main explanatory factors that have been identified for positive

transfer in TLA are (psycho)typology and proficiency in the given source language (Odlin & Jarvis, 2004; Rast, 2010; Ringbom & Jarvis, 2009). However, in addition to purely linguistic factors, it has also been suggested that metalinguistic awareness (MLA), commonly understood as "the ability to focus attention on language as an object in itself" (Jessner, 2006, p. 42), would explain why certain individuals are more inclined to produce positive transfer (Gibson & Hufeisen, 2003, 2006; Peyer et al., 2010). Indeed, increased levels of MLA seem to be characteristic of bi- and multilingual development (Bialystok, 2001a). Moreover, this "capacity to think abstractly about language" is believed to have facilitative effects on additional language learning (Bono & Stratilaki, 2009; Jessner, 2006; Kuile, Veldhuis, van Veen, & Wicherts, 2011; Rauch, Naumann, & Jude, 2012; Volgger, 2010). However, with regard to studies on positive transfer (see above), MLA has not been isolated as a variable. The goal of this research was to determine whether certain variables lead individuals to use their knowledge of English (L2) to uncover the meaning of lexical items in German (L3), and thus show a greater understanding of this additional language.

More specifically, the present study focused on a group of French-speaking Canadians and the ways in which they resorted to related vocabulary in their second language (L2), English, to understand new words in their third language (L3), German. In addition to the degree of similarity between the lexical items under investigation and their English translation equivalents, a number of learner variables were targeted in order to examine their relative contribution to this specific kind of "positive transfer" from a second to a third language. The main goal of the study was to examine the role of metalinguistic awareness (MLA) in the transfer process. The other factors under investigation were L2 proficiency and L2 exposure (including frequency/recency of use) as well as levels of interest in the target language and culture, German. The particular focus was on the cognitive dimension of crosslinguistic interaction. Even though researchers from different domains of cognitive science and applied linguistics appear to agree on the crucial role of MLA in the acquisition of first, second and third languages (e.g., Bialystok, Peets, & Moreno, 2014; Bono, 2011; Gombert, 1990), no study has, to our knowledge, examined this variable among the influential factors of positive transfer in a multilingual context.

This first chapter outlined the main inquiries that gave rise to our investigation. The second chapter will situate the study with respect to its theoretical underpinnings and to the central concepts underlying our research agenda. The third chapter provides a review of the literature on positive transfer, which will lead directly to our research questions. The fourth chapter will be devoted to the methodological approach that was adopted in order to arrive at a comprehensive account of the phenomenon in question, i.e., positive lexical transfer from English, and its potential predictors. Finally, the results of the study will be presented and briefly discussed in the fifth chapter, followed by chapter 6, which provides an in-depth discussion of findings with respect to each of the research questions, and examines the larger scope of the study, including its limitations and implications for future research. The seventh and last chapter aims at a comprehensive conclusion of our research endeavor, including a summary of the underlying goal of the study, its methodological approach and major findings as well as its possible implications and, finally, its contribution to the fields of second and third language acquisition.

2   Conceptual framework

This chapter provides the theoretical background and the conceptual framework that guide the present research. The concepts that constitute the heart of our study will be presented and analyzed with a special focus on speakers/learners of more than two languages. The first section (2.1) will be devoted to the notion of TLA and the theoretical orientations that distinguish this area of research from that of SLA. We will frame the specific approach within which we intend to examine L3 learners in the particular context of our study. Second, we will briefly look at the particularities of multilingual lexical development (2.2). In a third subsection (2.3), we will extend the discussion by focusing on metalinguistic development, and more precisely on the notion of *metalinguistic awareness* (MLA), which will be defined and analyzed with regard to the specific role it has been found to play in second and foreign language learning. Finally, the last subsection (2.4) will be devoted to the central issue of our study, which is that of *language transfer*, and the primary factors that appear to affect *lexical transfer* in TLA, specifically. Discussion of the preceding results will set the stage for the main study by introducing the implications for the investigation of *positive lexical transfer* in TLA.

2.1   Third Language Acquisition (TLA)

Learners of third and additional languages have long been considered on equal terms with second language learners. More precisely, research in second language acquisition (SLA) and bilingualism has tended to report on learners or users[2] of a second or third non-native language within the same theoretical boundaries established for L2 learners (Cenoz & Gorter, 2011, p. 357). Consequently, the term L2 could equally refer to any language beyond the first, a terminological solution which has restricted the view on subsequent languages in the learner's mind, since it does not account for the distinctive processes that learners of multiple languages engage in (e.g., De Angelis, 2007, pp. 4–7; Hammarberg,

---

[2] In bilingualism research, the focus has traditionally been on the sociolinguistic dimensions of being bilingual, rather than on the "process of becoming bilingual", the central focus of SLA (Cenoz & Gorter, 2011, p. 356). This explains the preference of the terms language *user* and *learner* in the respective fields.

2010, pp. 92–93). Throughout the last decade, a number of researchers have deplored this terminological and conceptual restriction, which has given way to the emergence of TLA as a field in its own right, as reflected by the increasing number of publications within this theoretical strand (Aronin & Hufeisen, 2009; Aronin & Singleton, 2012; Cenoz, Hufeisen, & Jessner, 2001; Cenoz, Jessner, & Hufeisen, 2003; De Angelis, 2007; Herdina & Jessner, 2002; Jessner, 2006; Ringbom, 2007; Singleton, Fishman, Aronin, & O Laoire, 2013).

Before we situate our own study within this field of research, we will give a brief outline of the theoretical orientations that have been present in SLA research to this day. An overview of these theories will illustrate the epistemological conceptions underlying our approach to learning in general and to TLA in particular.

2.1.1 Theoretical orientations in SLA research

Throughout the last three decades, a number of cognitively-oriented researchers have adopted a developmental perspective towards SLA, as pointed out by Hondo (2013). More precisely, in contrast to the generative principles of Universal Grammar (UG) (e.g., Chomsky, 1976), approaches such as Emergentism and Usage-based SLA, adopt a perspective on language learning based on interaction rather than on an innate capacity of abstract rule formation (see Hondo, 2013, p. 2). The opposition between generative and emergentist views of SLA has even been taken as an example of incompatible theories within the rationalist paradigm of cognitive linguistics. More precisely, Hulstijn argues that:

> …whereas generative linguistics and emergentism both allow abstract constructs not based on observable phenomena as long as testable hypotheses can be derived from the theory, the generative school explains language acquisition with the assumed existence of a language faculty, and the emergentist school assumes that language acquisition can be explained solely with a general cognitive learning device. (Hulstijn, 2014, p. 379)

Hulstijn (2014) points to the ontological assumptions, i.e., their fundamental postulates, which make the gap between nativist and developmental views unbridgeable. These reflections are part of a broader discussion about the epistemological embedding of SLA

theories and the possibility of a reconciliation of formerly thought incommensurable theoretical strands. This endeavor dates back to what has been labeled the cognitive-social debate of the 1990s (see Larsen-Freeman, 2007, for an overview). Namely, when researchers from the background of sociocultural theory entered the SLA stage in the mid-eighties (e.g., Frawley & Lantolf, 1985), their approaches did not seem compatible with the cognitivist tradition. A view of language acquisition in terms of a co-constructed process of individuals and groups of individuals in response to their environment implies different kinds of research questions and different methods of inquiry.

However, it seems crucial at this point to distinguish between methodology and epistemology. It is possible for a critical rationalist to use qualitative methods such as case studies or interviews, which in De Keyser's (2014) terms, are "hypothesis-generating" modes of observation, as opposed to quantitative methods that are usually associated with a rationalist approach. Mackey supports this claim:

> Simply associating cognitive with quantitative and social with qualitative is overly simplistic. […] Some of the constructs that we may consider socially informed, or at least useful for broadening the horizons of traditionally cognitive work – for example, motivation, willingness to communicate, or introspective measures – are psychological constructs and methods that have been included in research on language learning and teaching since the inception of the field. (Mackey, 2014, pp. 382–383)

Bearing in mind the possibility of what Mackey (2014) calls "partnerships of paradigms", it remains of crucial interest to consider Hulstijn's (2014) reflections on the incommensurability of certain theories:

> Differences in stance may be a matter of epistemology (extreme relativists vs. extreme rationalists), but they may alternatively constitute a matter of ontology (the part of the theory that is proposed as untestable). (Hulstijn, 2014, p. 380)

In other words, only when the ontological assumptions underlying a given theoretical strand are incompatible, is the gap between those strands unbridgeable. The developments discussed above reflect a diversity of theoretical approaches to SLA since its beginnings.

The cognitive-social debate of the 1990s has inspired a number of researchers to embrace socially-oriented methods of inquiry to complement their primarily cognitively-oriented perspectives (e.g., Larsen-Freeman, 1997, 2007; Swain, 2000; Tomasello, 2003). Even if some have spoken of a "social turn" (Ortega, 2011), this loosening of boundaries between strictly cognitive and strictly social investigations into SLA does not imply that rationalists have suddenly turned into relativists. Rather, a number of cognitivists acknowledge that both cognitive and social resources are relevant to language development, and that linguistic structures are transformed by use (e.g., Bybee, 2008). Accordingly, usage-based and emergentist approaches to SLA adopt data-driven methodologies, which explains a qualification of these strands as "exquisitely 'bottom-up'" (Robinson & Ellis, 2008, p. 511). Throughout the last decade, the gradual reconciliation of cognitive and social approaches to language learning has led to yet another theoretical dimension, in which the cognitive processes related to language learning are embedded in a constructivist perspective. What distinguishes this new theoretical strand commonly labeled Complexity Theory (Larsen-Freeman, 1997; 2012) from other developmental approaches to (cognitive) SLA is that it focuses on context in a broader sense, namely including physical, social and cultural dimensions, and, by extension, the learner's perspective on his/her own development. Based on the assumption that knowledge is constructed via social interaction, the excessively large number of interdependent variables make it impossible to predict what will be acquired by a specific learner and when.

In the light of what has been said about theoretical trends of more recent SLA theories, which have traditionally included any language beyond the first, the following sections provide an outline of research orientations with respect to multilingualism and the emergence of TLA.

2.1.2 Current perspectives on multilingualism

In a recent contribution on the typology of bi- and multilingualism, the notion of multilingualism has been qualified as "ambiguous" (El Euch, 2011, p. 58). In the present study, we will use the term with reference to individual multilingualism. In other words, we are interested in learners of multiple, in our case more than two, languages. It is important

to note, however, that within the boundaries of multilingualism research, it is often defined as a cover term to designate "the command and/or use of two or more languages by the respective speaker" (Herdina & Jessner, 2002, p. 52), and by extension, multilingual contexts as "those involving the acquisition of any language beyond a first or a second" (Odlin & Jarvis, 2004, pp. 123–124). Such a broad conception of multilingualism thus includes bilinguals[3], learners/users of second (see Trudgill, 2000) and of third or additional languages. It further presupposes that the acquisition and use of more than one language is the norm rather than the exception in the contemporary world (see Trudgill, 2000). In fact, research on multilingualism cannot be disconnected from sociolinguistic perspectives on the growing mobility of the world population, the spread of English around the world and the diversification of languages in use (Aronin & Singleton, 2008, pp. 4–5). In other words, the notion of multilingualism as a societal phenomenon can be seen as the starting point for a growing research interest in individual multilingualism from a psycholinguistic perspective (see authors cited in section 2.1).

The key concept in research on multilingualism, understood as including bilingualism of all types (Grosjean, 2013, pp. 7–11), is that of *multicompetence*, a term coined by Cook (1992), which refers to a "complex mental state including the L1 and L2 interlanguage, but excluding the L2"[4] (Cook, 2007, p. 17). Hence, instead of treating the learner/user of a non-native language as a deficient monolingual, the notion of multicompetence suggests a "unique form of language competence" that cannot be "measured […] against the yardstick of a native speaker" (Cenoz & Gorter, 2011, p. 357). Another important aspect of multiple language acquisition is the variability of interlanguage development, both with respect to the underlying knowledge of a given language (e.g., R. Ellis, 1985, p. 121) and to the actual use of that language in different contexts (Tarone, 1983, see also Grosjean's notion of language mode, 2001 and elsewhere). From a theoretical perspective, it is precisely the notion of variability that makes holistic approaches to multilingual development incompatible with the UG assumptions of the linearity of the acquisition process. In line

---

[3] The term "bilingual" in itself is ambiguous given that it can refer to different types of bilingualism, which is very rarely *balanced,* in the sense that both languages are used and mastered at similar levels and frequencies (Herdina & Jessner, 2002, p. 59)
[4] Cook uses the term L2 to refer to any language beyond the first.

with Cummins' (1991) conception of a *Common Underlying Proficiency* (CUP) which comprises "the cognitive/academic knowledge and abilities that underlie academic performance in both languages" (Cummins, 2005, p. 6), Cook (2002) argues that "the L2 user stands between two languages, even when apparently using only one, having the resources of both languages on tap whenever needed" (p. 5). In a multilingual perspective, this dynamic and thus unpredictable interaction between the psycholinguistic systems in the speaker's mind is thought to be the chief component of "multilingual proficiency" (Herdina & Jessner, 2002, pp. 57–58). As mentioned above, a growing research interest in the particular mechanisms of multiple language acquisition has led to the emergence of TLA as a distinct field. We will now give a brief outline of its theoretical foundations and of its basic conceptual considerations with respect to language learning.

2.1.2.1 Psycholinguistic approaches to multilingualism

Most of the researchers who have started to focus on the specific characteristics of TLA in the past fifteen years come from a psycholinguistic background. More precisely, their investigations have shown the importance of considering the distinctive aspects of TLA in order to account for complex processing mechanisms due to the interplay of more than two linguistic systems in the learner's mind (Aronin & Hufeisen, 2009; De Angelis, 2007; El Euch, 2011; Herdina & Jessner, 2002). On the structural level, TLA can be regarded as more complex than SLA, since learners have to discriminate between at least three languages, i.e., three or more systems of reference, in order to adapt to different communicative situations. At the same time, knowing more than two languages is also known to have facilitating effects on additional language learning, namely in terms of a multiplication of lexical and grammatical cues different languages may have to offer (e.g., Odlin & Jarvis, 2004; Ringbom, 2007). Even if the term L3 suggests a chronological view of language acquisition, it has been shown that the order of acquisition is only one of the numerous qualitative aspects identifying a particular foreign language in a learner's mind. According to Hammarberg (2010), factors like simultaneous acquisition, order-independent levels of proficiency, different types of language knowledge and different learning situations seem to contradict a chronological representation of languages (p. 94). Hence, the term L3 generally "refers to all languages beyond the L2 without giving preference to any

particular language" (De Angelis, 2007, p. 11).

Hufeisen's (2000) Factor Model (Figure 1) gives an extensive overview of the distinctive features of multiple language acquisition. Even though the main difference between TLA and SLA lies in the "foreign language specific factors" (in bold), both cognitive and affective factors are largely influenced by previous language learning experiences.
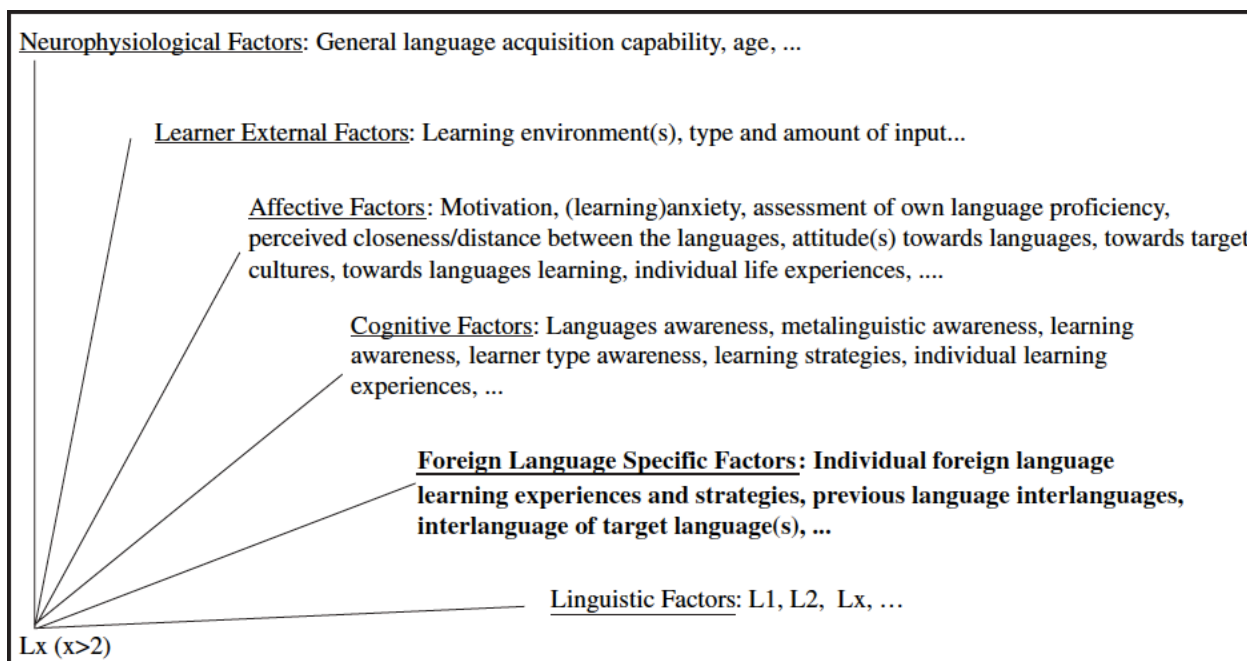
Figure 1. *Factors affecting multiple language acquisition (see Hufeisen, 2000)*

Even if social and contextual factors are taken into consideration by this model, its basic components are psycholinguistic in nature and have inspired researchers to investigate the development of a given Lx in individuals and groups of various language backgrounds (Cenoz, 2001; De Angelis & Selinker, 2001; Dewaele, 2001; Forsyth, 2014; Gabryś-Barker, 2006; Gibson & Hufeisen, 2003; Lindqvist, 2009; Llama, Cardoso, & Collins, 2010; Ó Laoire, 2006; Peyer et al., 2010; Ringbom, 2001).

2.1.2.2 A Dynamic Model of Multilingualism

The Dynamic Model of Multilingualism (Herdina & Jessner, 2002) has been developed in a psycholinguistic perspective with the specific goal to characterize the dynamics of

individual multilingualism. Inspired by Complexity Theory (see above), the authors utilize the theoretical basis of dynamic systems theory, especially known in meteorology, mathematics and psychology, to construct a model of multiple language acquisition, which is characterized by non-linearity, reversibility, stability and change of quality, among others (see Herdina & Jessner, p. 89). In this particular approach to multilingualism, language systems are analyzed against the theoretical background of dynamic biological systems whose main characteristic is their adaptation to the environment. In the case of language, the pressure of the environment is conceived of as a speaker's real and perceived communicative needs in a given context. The authors speak of a certain "language maintenance effort" (pp. 98-106), which, in case of a prolonged period in a monolingual environment, can lead to language attrition, even when a certain degree of language competence has been achieved in a language that is no longer used. In other words, both growth and decay are seen as natural processes in a dynamic system.

According to this approach, multiple language learning is a dynamic process depending on numerous factors related to each individual learner's history. Moreover, the constant interaction between the different languages of a multilingual generates new structures that are specific to each speaker and different from monolingual systems. In this sense, Herdina and Jessner's model is consistent with Cook's notion of multicompetence (see above). However, their conception of "multilingual proficiency" slightly differs from that of Cook's in the sense that it is derivable from the (assumed) levels of knowledge of each of the respective subsystems, instantiated by language use (p. 149). More precisely, their notion of multilingual proficiency is the function of two emergent properties: crosslinguistic interaction and a so-called "M(ultilingualism)-factor" (pp. 129-131). Whereas the former includes all types of transfer[5], as well as code-switching and borrowing, the latter refers to the synergetic effects of the cognitive development of multilinguals, particularly their presumably heightened levels of metalinguistic awareness[6], and an "enhanced multilingual monitor", which develops through the constant flux of activation and separation of different languages (p. 129). On the one hand, the M-factor

---

[5] The notion of *transfer* will be discussed in more detail in section 2.4.
[6] The notion of *metalinguistic awareness* will be discussed in section 2.3

represents the language management abilities of multilinguals, which might explain an accelerated development of language competence in a third or additional language (p. 130). On the other, possible interferences can slow down the acquisition process of an additional language system or lead to transformations in any other subsystem, including the native language(s).

Finally, what is specific to Herdina and Jessner's (2002) conceptualization of multilingual proficiency is that the three basic components – the speakers' competencies in each language subsystem, their language management abilities and crosslinguistic interaction – can compensate for each other in communicative situations (see p. 126). Therefore, even if a speaker lacks some relevant linguistic knowledge in a given context, his/her metalinguistic abilities can possibly make up for it. At the same time, crosslinguistic interaction can either be a source of disturbance or of facilitation for a given usage event. Accordingly, the authors argue that Chomsky's dichotomy between competence and performance is insufficient since it fails to account for the variability that has been observed in multilingual performance, especially since these deviations can be seen as "systematic […] in relation to the respective speaker systems" (p. 127). Instead of measuring performance data against the native speaker ideal, Herdina and Jessner suggest a focus on "communicative efficiency", which includes the knowledge of language resources and the ability to use these resources in various contexts (p. 128).

2.1.3 A fresh look at L3 learners

The present research adopts a psycholinguistic approach to the study of TLA. In line with Hufeisen's (2000) Factor Model (see Figure 1 above), our primary focus was on the foreign-language specific factors that seem to affect the acquisition of a third or additional language. More precisely, we investigated the role of a previously acquired language with respect to proficiency and exposure levels, as well as different usage-based variables related to individual learning experiences. In addition, we aimed at identifying the specific role of metalinguistic awareness in the acquisition process of an L3. Namely, in line with a dynamic view of multilingualism, we focused on the assets of experienced language learners, especially with respect to their metalinguistic development, which is thought to

affect further learning.

From the above discussion of current theoretical trends in SLA research and our alignment with holistic approaches to multilingualism (e.g., Cook, 2002), it follows that we adopt a developmental instead of a nativist view of language acquisition. More specifically, we believe that language learning is a situated process, which depends on the interaction between the learner and the environment. However, while inspired by the dynamic view of multilingualism, we do not adopt a constructivist approach, which would involve an investigation of the larger social context, as discussed with respect to complexity theory (see above). Rather, the aim of our study was to test the hypothesis whether metalinguistic awareness (MLA) facilitates language learning in a multilingual context. The central focus was on the positive influence of a typologically related background language on the comprehension of new vocabulary in the L3. Given that the precise goal was to identify the role of MLA in positive lexical transfer from L2 to L3, the following subsections will be devoted to these central concepts.

2.2   Lexical development in multilinguals

Naturally, the interaction of more than two linguistic systems in the learner's mind brings about a multiplication of lexical elements with different semantic and syntactic constraints as well as differing conceptual boundaries. The possibility of co-activation of different languages in a given communicative situation suggests that learners of two or more languages may have difficulties in keeping their languages apart. There is ample experimental evidence of learner errors resulting from crosslinguistic influence that interfere with the speaker's communicative goals (e.g., Paradis, 1993; Weinreich, 1953). Yet the question remains as to how efficient processing of lexical material is able to take place in the mind of a multilingual speaker. In the following subsection, we will discuss the difficulties entailed for multilinguals who need to find their way through a dense network of words from different languages. The second subsection will then provide an overview of the potential benefits that such a multiplication of lexical-conceptual structures in the mental lexicon may have for the language learner.

2.2.1 A complexification of the lexical network

Aiming at modeling lexical access in the multilingual lexicon, the notion of "language tag"[7] has proven particularly useful to explain how learners come to choose the right word from the right language to fit the context of communication. Thus, items from different languages may be competing for selection due to simultaneous activation based on semantic, syntactic and/or formal characteristics. The context would then provide the link to the appropriate item via the respective language tag attached to each lemma (Green, 1998). De Bot (2004) uses the notion of "language node" to refer to largely the same concept, in that it "controls the various processing components with respect to the language to be used" (p. 12). In sum, this context-dependent selection is also reflected in Grosjean's (1997, 2001) work on "language mode" which is mostly determined by the linguistic background of the interlocutor. Namely, a speaker of two languages is likely to switch freely from one to another when talking to another bilingual, given that there is no obvious constraint to comprehension. Dewaele (2001) has found the same tendency for multilingual speakers who adapt their amount of mixed utterances to fit the context.

However, the proposed existence of language tags does not explain instances where multilingual learners "fail[] to consciously recognize the source of [their] knowledge" when transferring lexical items from one language to another (De Angelis, 2005, pp. 10–11). Consequently, such a "system shift" would indicate that the appropriate tagging in the preverbal message has failed. Hence, De Angelis challenges the assumption that language tags are the key for appropriate lexical selection. If this were the case, she suggests that the respective tags may have "different strength values" that are expected to increase with proficiency and decrease when a language is not used (De Angelis, 2005, p. 14). In turn, "activation [of words] is about dynamically changing links in time" (de Bot & Lowie, 2010, p. 118). Therefore, given that lexical representations in the multilingual lexicon are inherently unstable and that "language tags are not firmly labeled" (Filatova, 2010, p. 93), the different dimensions of L3 vocabulary knowledge[8] are expected to be variable over

---

[7] In the field of SLA, this term was initially used by Green (1998).
[8] For example, Jarvis (2009) suggests five dimensions of word knowledge, namely (1) form, (2) meaning, (3) grammatical class and syntactic constraints, (4) collocations and syntagmatic associations, and finally (5)

time, even as the quantity of target language (TL) words in the mental lexicon increases (de Bot & Lowie, 2010; Meara, 2006; see also Milton, 2006, p. 246).

Apart from the context, which greatly influences language selection, de Bot (2004) argues that the respective levels of activation also depend on a number of other factors such as regular use, proficiency, age of acquisition, and method of instruction, amongst others. Furthermore, it is thought that each language of the multilingual has a default level of activation (de Bot, 2004, p. 10). Accordingly, those languages with higher levels of activation would be harder to deactivate. This would also entail that the native language always competes for activation given its high scores with respect to all of the factors that may lead to activation. However, numerous studies have shown the opposite, namely that multilinguals tend to resort to their more proficient L2 rather than to their mother tongue when communicating in their L3 (e.g., De Angelis, 2005; Singleton & Ó Laoire, 2006; Williams & Hammarberg, 1998). In order to explain this tendency with respect to activation patterns, de Bot (2004) suggests that the L1 may be more easily inhibited as a whole given the stronger links within its lexico-semantic system (p. 11). Moreover, this assumption could be seen as a partial explanation for the cognitive basis of the "talk foreign" phenomenon (De Angelis & Selinker, 2001, p. 56), discussed in section 2.4.2 under the label of *L2 status* as one of the variables affecting transfer between non-native languages (see below).

If we try to apply Green's model to multilingual contexts, the idea of inhibition and control implies that multilinguals need to put more effort into the selection process (e.g., Green, 1986, p. 215). This view has been adopted to account for so-called "neighborhood effects" in studies on word recognition, where the existence of phonetically similar items across languages led to slower reaction times, which in turn reflected the interference of non-target words with the selected TL due to crosslinguistic activation (Van Heuven, Dijkstra, & Grainger, 1998). It should be stated, however, that this 'cost of switching' has mostly been documented in studies on bilingual learners (e.g., Meuter & Allport, 1999). Interestingly,

---

lexical and conceptual associations. In Nation's (2001) view, two additional dimensions should also be considered, namely (6) frequency of use, and (7) register.

while these findings indicate that more competing items might slow down the selection of the appropriate candidate in word recognition, contrasting findings have been reported in Costa and colleagues (2006) with respect to speech production. Namely in discussing data from neighborhood and cognate studies, they conclude that reaction times tend to be shorter if there are many neighbors or cognates from other languages known to the speaker (Costa, La Heij, & Navarrete, 2006, pp. 143–146). Instead of assuming that the multiplication of lexical material from different languages leads to confusion, findings like these actually indicate that knowledge of formally and/or semantically related items from other languages can also help learners to process L3 vocabulary more efficiently.

2.2.2 A facilitation for further language learning

Based on a number of large-scale studies conducted in Finland, Ringbom (2007) concluded that learners tend to rely mostly on formal aspects of new words in order to understand TL items, especially at initial stages of vocabulary acquisition, given that "formal correspondences arouse hopes of semantic or functional equivalence" (p. 9). More specifically, in the Finnish context, it was the speaker's prior knowledge of Swedish which helped Finnish-Swedish bilinguals to process English L3 vocabulary. This tendency to interpret L3 vocabulary via the L2 lexicon was mainly explained on the basis of typological proximity between Swedish and English, whereas Finnish couldn't provide many lexical cues for this particular group of L3 learners. This tendency to resort to a typologically related source language for both comprehension and production of a given L3 has been confirmed in numerous other studies investigating lexical access in the multilingual lexicon (e.g., Cenoz, 2003b; Gibson & Hufeisen, 2003; Odlin & Jarvis, 2004; Rast, 2010). Consequently, the proposed "lack of proper semantic organization for foreign language words" (Meara, 2009, p. 17) may also be regarded as one of the major reasons for foreign language learners to make use of what has been referred to as "potential vocabulary" (Ringbom, 1992, p. 88). In sum, learners of typologically related languages are expected to make more appropriate form-meaning connections in their L3 than learners with entirely unrelated background languages, given that close typological relationships often imply similarities between many sub-features of the lexicon (Wesche & Paribakht, 2010, p. 34).

Moreover, it is not only cognates that may allow an advantage for the acquisition of subsequent TL vocabulary. As has been argued with respect to growing vocabularies of language learners, the mere fact of knowing existing synonyms of a given target item seems to help learners make faster form-meaning connections, i.e., learn the words, as reflected in the tasks evaluating both receptive and productive vocabulary use (Webb, 2007). If we consider languages as "systems of synonyms" (Filatova, 2010, p. 86), L3 vocabulary would also be acquired more efficiently given the multiplication of known synonyms from other languages, which provide lexical cues. This tendency has also been observed with respect to lexical inferencing in foreign language reading: "In such cases, even a partially appropriate lemma will facilitate comprehension" (Wesche & Paribakht, 2010, p. 20). However, this view may be simplified, since conceptual representations are not always shared across languages (see above). Still, learners who have already experienced the need to restructure concepts in a second language would be likely to display a greater flexibility with respect to potential differences in the structure of L3 concepts. This assumption is closely related to the cognitive dimensions of TLA (see following section).

2.2.3 Interim summary: Implications for L3 vocabulary acquisition

Investigations in foreign language vocabulary acquisition have led researchers to diverse approaches concerning the organization and the development of mental lexicons of language learners in comparison to native speakers. As far as L3 learning is concerned, the multiplication of words from different languages has already been interpreted as a hindrance, given that speakers are not only faced with a multitude of new linguistic patterns to handle but that they also need enhanced control mechanisms to keep their languages apart (e.g., Green, 1998). Yet the abundance of lexical material in the mind of a multilingual may also have facilitative effects on the learning process of new words in a foreign language, namely due to the presence of *potential vocabulary*, that is words from other languages that bear formal and functional similarities with TL items (e.g., Ringbom, 2007). This perspective is also reflected in the notion of "learning burden" which is supposed to be lighter if "a word represents patterns and knowledge that the learners are already familiar with" (Nation, 2001, pp. 23–24). Finally, speakers who have already learned a foreign language, are not only expected to have enhanced metalinguistic abilities,

which should help them for subsequent learning (see next section), but they are also experienced with respect to learning strategies that might be of use for the retention and acquisition of new TL vocabulary. In the context of the present study, the focus will be on one particular vocabulary learning strategy that involves the recognition of related background vocabulary from a previously acquired language for the comprehension of unknown words in an additional language. It is their immediate reaction to the target material throughout this first encounter that will be in the center of investigation. According to the above-stated observations, participants are expected to activate lexemes from their other languages, especially from the typologically related L2, English, in order to make sense of new L3 vocabulary.

2.3   Metalinguistic development in multilinguals

If we assume that learners naturally rely on already existing linguistic and world knowledge when learning a foreign language, then structural analogies between the languages known to a learner would be expected to have a positive influence on the learning process (e.g., Ringbom & Jarvis, 2009). However, learners seem to make use of their prior knowledge to different extents. As suggested by Hufeisen's (2000) Factor Model (see above), individual factors such as learner type, language learning experience and levels of metalinguistic awareness (MLA) are probably, at least partly, responsible for such discrepancies in the use of crosslinguistic similarities as a learning strategy. More precisely, it seems as though the cognitive development of more experienced language learners facilitates subsequent language learning (e.g., Bialystok, 2001b). This section aims at providing a comprehensive picture of the notion of metalinguistic awareness and its relationship to L2 (or L3) proficiency as well as the effects that MLA is thought to have on foreign language learning. In order to examine the specific role of MLA in learning a foreign language, the next subsection will be devoted to defining the concept of MLA, taking into consideration different theoretical approaches depending on the respective theoretical background. Only then will it be possible to discuss the distinct role of MLA both as a potential source and as an outcome of L2 (or L3) learning. While MLA had initially been studied especially with respect to the cognitive development of bilingual children, we will mainly focus on more recent findings in adult foreign language learning, to examine the role of MLA and its

presumably positive effects on the learning process.

2.3.1 Definitions

In order to define the notion of "metalinguistic awareness", it will be necessary to comprehend each of its components. We will first provide a brief outline of "metalanguage" on the basis of which the adjective "metalinguistic" has come to be used in both linguistics and psycholinguistics, yet with somewhat different research orientations. The second subsection will then serve as a short introduction to "awareness", such as it has been treated in SLA research. Only then will we focus on "metalinguistic awareness" and the terminological issues that have been raised in the literature. Thereafter, MLA will be differentiated from "metalinguistic abilities", "metalinguistic knowledge" and "metalinguistic tasks", since all of these terms will be used in the main discussion on the role(s) of MLA in the development of foreign language proficiency.

2.3.1.1 Metalanguage

Broadly speaking, when language becomes its own content, we speak of *metalanguage*. The scholarly interest in this "in-built self-referential potential of linguistic systems" (Coupland & Jaworski, 2004, p. 17) goes back to Russian Formalism. Roman Jakobson (1960) shaped the term *metalanguage* in his most influential writings on the *functions* of language half a century ago. From a linguist's point of view, "metalanguage" strictly refers to words that are used to describe words and structures in their functional sense. Thus, any verbal reflection about the functional properties of language, but also any linguistic categorization such as the identification of word class, for instance, reflects metalanguage. From a psychological perspective, on the other hand, the term "metalinguistic" refers to the human abilities, behaviors, attitudes or processes related to metalanguage (Pinto, Titone, & Trusso, 1999, p. 13). Situated at the crossroads of these two disciplines, the field of developmental psycholinguistics has seen the emergence of a theoretical framework around the notion of 'metalinguistic development' from the 1980s onwards, especially with the contribution of an Australian group of researchers (Tunmer, Pratt, & Herriman, 1984) and later those of Gombert (1990, 2000), amongst others. As opposed to several authors who have identified young children's behavior when correcting their own utterances as

"metalinguistic" (e.g., Clark & Andersen, 1979; Clark, 1978; Marshall & Morton, 1978, as cited in Pinto et al., 1999, p. 18), an important distinction has been made between such rather spontaneous forms of self-correction and the conscious control of the structural characteristics of language. With reference to his earlier work, Gombert (2000) insists that the latter does not merely reflect a higher degree of development but that it represents a qualitative difference, which ought to be reflected in the terminology:

> Pour éviter les assimilations déformantes, une distinction doit être faite entre les capacités manifestées dans les comportements spontanés (par exemple, le trouble engendré chez le jeune enfant par l'audition d'une phrase agrammaticale) d'une part et, de l'autre, les capacités qui sont fondées sur des connaissances mentalisées et intentionnellement appliquées (par exemple, corriger la syntaxe d'un texte écrit). C'est plus qu'une différence de degré qui sépare ces deux ensembles de comportement, mais une différence qualitative dans les activités cognitives elles-mêmes. Pour des raisons de clarté terminologique, il convient donc d'éviter d'utiliser le même terme pour qualifier ces deux types de comportements. (Gombert, 2000, p. 1)

Gombert (2000) further suggests a terminological distinction based on Culioli's (1968) notion of "epilinguistic", which encompasses all processes where the linguistic knowledge of a subject intervenes in the language he or she uses, and becomes manifest on the surface level, such as in self-corrections.

> Posant comme principe qu'un caractère réfléchi et intentionnel est inhérent à l'activité strictement métalinguistique, il [Gombert, 1999] propose d'utiliser le terme "épilinguistique" pour désigner les comportements qui, bien qu'isomorphes aux comportements métalinguistiques, ne sont pas le résultat d'un contrôle conscient par le sujet de ses propres traitements linguistiques. (Gombert, 2000, p. 1)

Others have identified such processes as "intuitive awareness" (e.g., Tunmer & Herriman, 1984, p. 19) or "intuitive knowledge" (Alderson, Clapham, & Steel, 1997, p. 95). However, these processes can be made explicit through conscious manipulation. It is thus the passage towards deliberately applied knowledge that identifies "metalinguistic" activity. In other words, the distinction between epilinguistic and metalinguistic processes seems to be based

on the intentional dimension of consciousness, rather than on depth of analysis. Hence, it would be possible to speak of metalinguistic activity if a person is consciously and deliberately manipulating structural elements of a language, yet without necessarily implying that he or she is aware of the (explicit) grammatical rules underlying the given structures. On the basis of these terminological observations regarding its basic constituents, we believe that MLA *can*, but does not necessarily, imply the explicit knowledge of underlying linguistic structures, such as reflected in the appropriate use of terms for grammatical categories and rules.

2.3.1.2 Awareness

The first distinction to be drawn in order to define the concept of awareness is that between "explicit" and "implicit" learning and knowledge, respectively. According to Schmidt (1994), explicit learning requires awareness, whereas implicit learning does not. And "awareness", for Schmidt, is the learner's knowledge – in terms of a subjective experience – that he or she is detecting a stimulus. However, this does not necessarily result in *explicit knowledge* (a product), one that can be accessed on demand, as for instance, by naming grammatical categories. Explicit *learning* tends to be used interchangeably with explicit *knowledge*. However, Schmidt (2001, 2010) further underlines the distinction between "noticing", the conscious detection of a stimulus, and "understanding", the recognition of a general principle, rule or pattern. Whereas *noticing* is limited to the surface structure of an utterance, *understanding* represents a deeper level of awareness, since it entails the explicit knowledge of underlying rules (e.g., Schmidt, 2010, p. 725). Hence, explicit learning – i.e., learning with awareness – *can* result in explicit knowledge, which is highly analyzed and thus permits extrapolation of underlying rules. But at a lower stage of awareness – i.e., *noticing* without *understanding* – the result is less analyzed knowledge and not necessarily generalizable as a grammatical rule. In sum, awareness in language learning has often been equated with explicit knowledge of grammar, whereas the initial stage of awareness is merely reflected in the conscious detection of a stimulus, at the level of *noticing*.

Naturally, the study of awareness requires the definition of certain indicators providing evidence of a conscious learning experience. In fact, not only must a change in verbal

behavior be observable such as the appropriate use of the TL structure after extended exposure to it. But the learner also has to report that he or she was aware of this experience as it took place, thus reflecting a kind of "meta-awareness" (Leow, 2000, p. 560). Finally, the learner should then be able to describe the experience, namely in articulating the underlying rule(s) governing the structure that has been acquired. According to Leow (2000), the latter case reflects "high" awareness, whereas merely reporting the learning experience without identifying the underlying structure of the given feature would be indicative of "low" awareness (p. 560). In a more recent contribution, he suggests three levels of awareness, the first of which is *noticing*, followed by *reporting* and finally *understanding* (Leow, Johnson, & Zarate-Sandez, 2011, p. 70).

As will become clear in the last subsection, metalinguistic tasks generally reflect these levels yet mostly without a focus on understanding in terms of articulating the underlying structure of linguistic features, using appropriate metalinguistic terminology (see below).

2.3.1.3 Metalinguistic Awareness (MLA)

It was mainly with respect to the cognitive development of children that psycholinguists began to investigate MLA, thus referring to "the ability to reflect upon and manipulate the structural features of spoken language, treating language itself as an object of thought, as opposed to simply using the language system to comprehend and produce sentences" (Tunmer & Herriman, 1984, p. 12). More recent contributions in the fields of second and third language acquisition, which investigated the verbal behavior of adults learning foreign languages, have provided largely equivalent definitions of MLA, namely in terms of the "ability to focus attention on language as an object in itself or to think abstractly about language, and, consequently, to play with or manipulate language" (Jessner, 2006, p. 42, see also El Euch, 2010, p. 18). However, as reflected in both the above-stated quotations where "awareness" appears to equal "ability", several researchers interested in metalanguage as an indicator of cognitive development have addressed this terminological inconsistency. In particular, Pinto, Titone and Trusso (1999) noted that "metalinguistic awareness" along with notions such as "language awareness" and "declarative knowledge of the rules of language" has been used to refer to "[*a*] *particular ability* (for example, the

ability to identify a grammatical error, specifying the general rule which applies)" (p. 35, italics and parenthesis in the original).

In order to tease apart the different concepts behind these terms, we will first draw on a more straightforward definition of MLA, which will allow for a distinction between metalinguistic *abilities*, metalinguistic *knowledge*, and metalinguistic *tasks*, respectively, as presented in the following subsection. According to Bialystok (2001a), "[m]etalinguistic awareness implies that attention is actively focused on the domain of knowledge that describes the explicit properties of language." She further refers to it as "a momentary phenomenon, something achieved at a point in real time because attention has been focused on certain mental representations" (Bialystok, 2001a, pp. 126–128).

2.3.1.4 Metalinguistic abilities and metalinguistic knowledge

As reflected in Bialystok's (1993, 2001) work, metalinguistic *awareness* represents the conscious control over the components of linguistic structures in use. Both from a conceptual and from a methodological perspective, it is crucial to consider the different applications of the term "metalinguistic", in order to arrive at a working definition of MLA.

> When applied to tasks, certain uses of language, such as making repairs or judgments about well-formedness, are classified as metalinguistic or not. A learner performing a task classified as metalinguistic is demonstrating metalinguistic ability. […] When applied to levels of awareness, virtually any performance can have metalinguistic properties if it is carried out with the deliberate control and awareness of the language learner. (Bialystok, 1993, p. 217)

This distinction makes it clear why metalinguistic *awareness* and *ability* have often been used interchangeably, since the latter is included in the former (Pinto et al., 1999, p. 36). However, the notion of "awareness" can either refer to a momentary activity or to a latent state. Whereas Bialystok (2001a) insists on the momentary nature of MLA, Jessner (2006) refers to it in terms of a potential, i.e., "the ability to focus attention on language as an object in itself" (p. 43). Naturally, one must have this ability in order to become (momentarily) aware. But having such abilities does not necessarily entail that the person will always be aware. For reasons of clarity, a terminological distinction between

awareness and ability was judegd to be useful, namely because MLA manifests itself as a momentary phenomenon, i.e., an activity. Finally, we decided to define MLA as follows: *Someone who is metalinguistically aware has the ability to deliberately direct his or her attention to the structural features of language, and thus to manipulate its components.*

A further distinction should be drawn between MLA and metalinguistic knowledge. Typically, the latter term is used to refer to the "explicit knowledge" about the given language. More specifically, Alderson, Clapham and Steel (1997) suggest that "whatever the explicit knowledge looks like, it must include metalanguage, and this metalanguage must include words for grammatical categories and functions" (p. 97). Accordingly, the concept of metalinguistic knowledge has mostly been treated in terms of knowledge about the grammar of the target language (e.g., Elder & Manwaring, 2004; Hu, 2011; Roehr & Gánem-Gutiérrez, 2009; Serrano, 2011). However, as stated above, our notion of MLA does not necessarily entail the correct use of grammatical categories to describe the linguistic features being manipulated. This distinction is important with respect to the type of measurement that will be used to assess MLA in the present study.

2.3.1.5 Metalinguistic tasks

In practice, MLA is reflected through linguistic activities, which, in turn, can be measured by means of metalinguistic *tasks*. However, a variety of tasks have been used within and across research paradigms, which reflects, once again, the complexity of the notion of MLA that is determined or somehow redefined by the nature of the metalinguistic task chosen for the specific context of each study. Within a cognitivist framework, Bialystok identified two major constituents of MLA, namely the cognitive processes of "analysis of representational structures" and "control of selective attention" (Bialystok, 2001b, p. 177, based on earlier work from 1993). Along these lines, Lefrançois (2000) makes a distinction between tasks that call for more *analysis* (e.g., grammaticality judgment, identification of ungrammatical elements, correction of such elements, explanation of the correction) and for more *control* (e.g., word substitution ignoring the meaning, identification of rhymes and synonyms of the target item). According to her, tasks to elicit data on MLA need to be varied, some focusing more on analysis, others more on control (p. 71). A number of other researchers in the

fields of educational psycholinguistics (e.g., Pinto et al., 1999) and SLA/TLA (e.g., Gibson & Hufeisen, 2006; Lasagabaster, 2001; Rauch et al., 2012; Renou, 2001; Serrano, 2011) have also drawn upon the analysis/control distinction in their conceptual frameworks and the respective methods of data collection have proven to be quite varied, both across and within each of the respective paradigms. According to Renou (2001) a reliable measure of MLA should "require learners to access and elaborate upon their linguistic knowledge", since the task results would otherwise merely reflect knowledge of grammar (p. 251). In other words, what should be reflected in a metalinguistic task is *how* implicit knowledge turns into explicit knowledge (see El Euch, 2010, p. 18). An interesting observation concerning the analysis/control distinction has been made by Rauch and her colleagues (2012) who – with reference to Bialystok's work – stress the fact that measures of control do not seem to vary a lot among bilinguals (or L2 learners) at different levels of L2 proficiency, whereas the analysis component of MLA appears to increase with higher levels of proficiency (p. 406). Accordingly, their measure of MLA was solely based on this second component, given that the focus of their study was on the role of MLA in relation to proficiency levels in L1, L2 and L3. Even if researchers do not always elaborate on their choice of metalinguistic task with respect to Bialystok's model of MLA, there is clearly a preference for tasks that focus on "analysis of representational structures", which may be explained by Rauch's (2012) rationale.

The most frequently used tasks are grammaticality judgments and error correction tasks, usually in the L2 or Lx, which have also been used in combination (e.g., Alderson et al., 1997; Gibson & Hufeisen, 2006; Lefrançois, 2000). Moreover, many of these are followed by explanation or justification tasks, essentially with the goal of "test[ing] the learners' ability to implement pedagogical grammar rules" (Roehr, 2007, p. 180). Elder and Manwaring (2004), for instance, assessed "grammatical knowledge" (p. 148) by means of a test that required their participants to match grammatical categories to parts of speech and to correct and explain errors (pp. 149-150). Likewise, Roehr and Gánem-Gutiérrez (2009) operationalized the construct in terms of "L2 description/explanation ability and L2 language-analytic ability" (p. 175), using a comparable protocol. As stated above, a similar focus on the capacity of identifying and naming grammatical categories and rules is present

in other studies investigating "metalinguistic knowledge" (e.g., Alderson et al., 1997; Hu, 2011). As outlined above, we do not adhere to this restricted notion of MLA for two reasons. First, the appropriate use and manipulation of grammatical categories and rules, which determines performance on the tasks stated above, represents a specific type of highly analyzed knowledge. To fail these tasks can, however, not be taken to reflect absence of metalinguistic knowledge (or ability) altogether. Second, if the focus is on awareness rather than on specific representations of linguistic knowledge, the following considerations on the manifestations of MLA should be taken into account:

> It is not sufficient merely to perform a metalinguistic task (such as repair), nor to have demonstrated an independent ability to focus on language forms, but performance must indicate as well that the learner is in a particular mental state, namely that the learner is aware of the forms and functions of the language being manipulated. (Bialystok, 1993, p. 217)

Furthermore, we support the claim that the performance on any metalinguistic task carried out in the learner's L2 or Lx will, at least partly, depend on proficiency levels in the respective language (e.g., Kuile et al., 2011; Rauch et al., 2012). To our knowledge, the most comprehensive tool for a systematic measurement of MLA that allows for a classification into different levels is the Metalinguistic Ability Test (MAT)[9] developed by Pinto and her colleagues (1999). Even though this test battery has been used by other researchers in subsequent studies on the role of MLA in foreign language learning (e.g., El Euch, 2010; Lasagabaster, 2001; Pinto, Trusso, & Kristiansen, 2002), studies with a focus on multiple language learning tend towards intro- and retrospective methods such as think-aloud protocols, questionnaires or interviews to elicit data on metalinguistic processing (e.g., Gibson & Hufeisen, 2003; Jessner, 1999, 2006; Volgger, 2010).[10]

Obviously, the question is whether these differences with respect to the definition of MLA and to the research designs make it possible to determine general tendencies as to the relationship between MLA and foreign language proficiency, given that different

---

[9] A detailed description of the instrument will follow in the corresponding section of the Methodology (Chapter 4).
[10] We will elaborate on these studies in Chapter 3.

researchers may not be measuring the same thing. This question will be addressed in the following section.

2.3.2 The relationship between MLA and proficiency in L2 or L3

First of all, MLA has commonly been treated as a consequence of bilingualism, yet not necessarily as dependent on a high level of proficiency in a given L2 or L3.[11] We will first give an outline of the role of foreign language knowledge for cognitive development, as it has been reflected in studies in bilingualism and SLA. In the second subsection, we will extend the discussion to multilingualism and elaborate on the role that MLA has been attributed in the study of TLA, namely focusing on speakers of more than two languages. It will thus become clear how the above-stated research designs have led to slightly different observations, yet indicating the same tendency, namely pointing towards a crucial role for MLA in the process of learning a second, third or additional language.

2.3.2.1 Bilingualism and SLA research

In line with findings on the relation of language development and cognitive skills in the field of developmental psychology (e.g., Piaget, 1929; Vygotsky, 1962), Bialystok's work on the cognitive development of bilingual children has confirmed that the manipulation of two linguistic systems provides them with a natural understanding of the arbitrary relation between words and their real-world referents (Bialystok, 2001b, p. 171). On the basis of this level of abstraction, speakers of two languages seem to display an "advantage in cognitive processing" as reflected in a number of metalinguistic tasks (p. 180). Comparable observations have also been made with adult speakers of two or more languages (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004).

In the field of applied linguistics, it was Peal and Lambert's (1962) study, in which the relation between bilingualism and cognitive development[12] was first made explicit. These

---

[11] This aspect has been addressed above, when introducing the notion of "common underlying proficiency" (Cummins, 1991), but will be further discussed below.

[12] Levels of "intelligence" were measured by means of the *Lavoie-Laurendau* test, the *Raven Progressive Matrices Test* as well as the *Thurstone Primary Mental Abilities Test*, all of which are not directly concerned with metalinguistic processing, but rather with cognitive abilities, in general.

researchers attributed the relative success[13] of bilingual learners over their monolingual peers to higher levels of cognitive flexibility. Ever since, the potential cognitive advantages related to foreign language learning have received more attention in both SLA and TLA research (Jessner, 2008a, p. 277). Thus, the individual speaker's level of MLA is not only thought to increase in the course of learning a foreign language, but it is also expected to accelerate the learning process of subsequent languages (e.g., Bialystok, 2001a; Dillon, 2009; Jessner, 2006; Lasagabaster, 2001; Pinto et al., 2002; Ransdell, Barbier, & Niit, 2006).

Throughout the last decades, a number of researchers have further investigated the relationship between MLA and different types of proficiency measures in the given target language (e.g., Alderson et al., 1997; Elder & Manwaring, 2004; Ranta, 2002; Rauch et al., 2012; Roehr, 2007). These investigations have yielded somewhat different results with respect to this relationship. As we have hinted at above, the fact that Roehr (2007) found significant correlations between metalinguistic knowledge and L2 proficiency measures might, at least partly, be due to the fact that metalinguistic skills were measured in the L2. However, Elder and Manwaring (2004), whose metalinguistic tasks consisted of the same components as Roehr's (see above), only found weak or no correlations with L2 performance. The authors suggest that this finding might be related to the type of instruction that the different groups of participants had received (p. 159). Furthermore, when there was a relationship to be established, it was stronger with the skills of L2 reading and writing than with speaking and listening, which, in turn, is consistent with the basic assumption that literacy is related to metalinguistic development in the L1 (e.g., Gombert, 1990; Ryan & Ledger, 1984) and in the L2 or L3 (e.g., Rauch et al., 2012). Another large-scale study using metalinguistic tasks that assessed knowledge of grammar (Alderson et al. 1997) revealed very weak correlations with L2 proficiency measures. Moreover, the authors found that the part of the metalinguistic task which consisted in naming grammatical categories was not correlated with the error correction nor with rule extrapolation that did not necessitate metalanguage (see pp. 106-108). Similarly, the results of Elder and Manwaring's study (2004) revealed "the ability to correct erroneous sentences

---

[13] Linguistic abilities were measured by means of a word association, a word detection and a vocabulary test.

did not always go hand in hand with the ability to explain the rule which had been violated in that sentence" (p. 156). In line with our conception of metalinguistic awareness, this suggests that there seems indeed to be an intermediate level of awareness, where learners are able to reflect on and manipulate language without necessarily stating explicit rules or grammatical categories as such (see Leow et al., 2011).

2.3.2.2 Multilingualism and TLA research

As pointed out in the first section of this chapter (2.1), holistic approaches to bi- and multilingualism (e.g., Cook, 2002; Grosjean, 2001) and the emergence of the field of TLA have drawn more attention to the role of MLA in the learning of second and third languages. Nearly all studies with a focus on the interaction of linguistic and cognitive resources in TLA argue for a crucial role of MLA in L3 development (e.g., Bono & Stratilaki, 2009; Cenoz, 2003a; Gabryś-Barker, 2006; Hufeisen & Gibson, 2003; Peyer et al., 2010; Singleton & Aronin, 2007). The ones that have addressed the facilitative role that MLA is thought to play in positive transfer between second and third languages (e.g., Jessner, 1999; Gibson & Hufeisen, 2003; Dillon, 2009) will be discussed in the literature review (Chapter 3). Even if few studies explicitly measure MLA in relation to proficiency, the observable tendency supports the claim for a reciprocal effect between MLA and foreign language learning. For instance, the results obtained by Rauch and her colleagues suggest that MLA mediates the positive effect of biliteracy on L3 reading (Rauch et al, 2012, p. 411). In this study, MLA was measured via the so-called Language Awareness Test (Fehling, 2008), which required learners to apply word building rules from English (L3) to Swedish and Dutch, both unknown to the German-Turkish bilinguals who participated in the study. Stepwise multiple regressions revealed that full biliteracy predicts MLA, while MLA predicts L3 reading skills. Another recent study where MLA was operationalized in terms of an inductive learning ability via the extrapolation of rules from an unknown language (Kuile et al., 2011) revealed that learners with a background of bilingual education (Dutch and English) scored higher on MLA than those who attended a Dutch only school. Moreover, higher scores on the MLA measure correlated with higher L2 proficiency and with the number of languages spoken at home (Kuile et al., 2011, p. 239). These results further strengthen the claim that the mastery of more than one language

facilitates the acquisition of subsequent languages via the development of enhanced metalinguistic abilities, which, in turn, seem to be fostered by learning experiences in educational contexts.

In this perspective, instead of being viewed as a possible side effect, MLA is taken to represent a central driving force of multilingual development (Herdina & Jessner, 2002, p. 116). Moreover, it is a concept that has commonly been treated as language-independent, namely with reference to the notion of *Common Underlying Proficiency* (see above), and thus appears to be applicable to any one of the speaker's languages. Indeed, the vision of MLA as a constituent of a presumed CUP has also been widely adopted by researchers focusing on third and additional language acquisition (e.g., Abu-Rabia & Sanitsky, 2010; Dillon, 2009; El Euch, 2010; Jessner, 1999, 2006). In this sense, *metalinguistic awareness* is taken to reflect the implicit metalinguistic knowledge made explicit via metalinguistic tasks (see above) in a given L1 or Lx.

Finally, even though several researchers have pointed to the fact that MLA is not easily operationalized, they unanimously state the "catalytic effect" it is supposed to have for the language learning process (Herdina & Jessner, 2002, p. 116). In a recent study on the multilingual's potential use of "affordances", Singleton and Aronin (2007) support this claim in stating that the condition of learning more than one foreign language is particularly favorable to an increase in language awareness, which in turn should facilitate subsequent learning. According to them, *affordances* are intimately related to awareness, given that "the higher the level of language awareness, the more effectively language-related possibilities are likely to be perceived…" (p. 85).

2.3.3   Interim summary: The role of MLA in TLA

As discussed in this subsection, *metalinguistic awareness* is a complex concept. The variety of methods that have been used to elicit data on MLA does not make it a simple task to determine its role for foreign language learning. Even though bilinguals have been shown to display more cognitive flexibility with respect to metalinguistic tasks than monolinguals (e.g., Bialystok, 2001a; Peal & Lambert, 1962; Pinto et al., 2002) the assumption that

multilinguals, i.e., speakers of more than two languages, should reach even higher levels of MLA, has not yet been confirmed. Furthermore, a lack of precision in the identification of MLA still seems to prevent researchers from providing evidence for such increased levels of awareness in speakers of more than two languages. The following statement reflects the situation fairly well:

> [G]iven […] that awareness is a psychological state that cannot be measured with any precision, it would be difficult to demonstrate the veracity of a sweeping statement such as that multilinguals are always more aware of their language resources and of the opportunities available to them. However, it would be extraordinary if the multilingual experience did not lead – at least in a large proportion of instances – to increased levels of such awareness. (Singleton & Aronin, 2007, p. 85)

Speakers who use their metalinguistic abilities to draw conclusions based on crosslinguistic similarities and contrasts will easily decode new structures, be it syntactic, semantic or phonological constraints of the given TL (e.g., Gibson & Hufeisen, 2003; Jessner, 2006; Peyer et al., 2010; Singleton & Aronin, 2007) or entirely new scripts altogether (e.g., Abu-Rabia & Sanitsky, 2010). However, it doesn't seem unconditional that higher levels of MLA lead to accelerated learning. The metalinguistic abilities that a language learner may have at his or her disposal must be intentionally focused on the given TL in order to have a positive effect on the learning process.

2.4   Language transfer in TLA

Prior to discussing the most important factors that seem to affect language transfer in general and lexical transfer in particular, in both L2 and L3 learning, we will first provide comprehensive definitions of the relevant terminology. More precisely, a concise overview of *language transfer* as a cover term will help define the second and central notion of this discussion, namely *lexical transfer*, including its subcategories as identified in the literature. We will then delineate the influential factors affecting the interaction between non-native languages, with a focus on learners of three or more languages.

2.4.1 Definitions of language transfer

Generally speaking, the notion of *language transfer* can summed up as the "influence resulting from similarities and differences between a target language and any other language that has been previously (and perhaps imperfectly) acquired" (Odlin, 1989, p. 27). Depending on a number of language and learner-related factors such as, for instance, language proficiency and typology, prior language knowledge will have more or less influence on the learner's comprehension and production of the TL. In other words, elements from a "source" language are being transposed to a "recipient" language (Jarvis & Pavlenko, 2008, p. 22), which is, in most investigations, also the TL, i.e., the one that is currently being acquired. In addition, researchers have also investigated instances where later acquired languages influence the learner's L1 or an earlier acquired L2 (e.g., Ecke, 2008; Jarvis, 2003). Technically speaking, transfer processes may take place between all language systems known to the learner, and in any direction (Jarvis & Pavlenko, 2008, pp. 21–22). Furthermore, the interaction between different language systems in a learner's mind may either help them understand or produce the TL or not. Broadly speaking, such processes that hinder comprehension (e.g., false cognates) or result in production errors or language switches, are commonly referred to as "negative transfer" (Selinker, 1969, p. 91), while those instances where another language provides cues that facilitate the comprehension or production of TL elements are referred to as "positive transfer" (Jarvis & Pavlenko, 2008, p. 25; Odlin, 2003, p. 438). Traditionally, the study of language transfer was mainly concerned with "interferences" from the mother tongue (Corder, 1979; Lado, 1957; Selinker, 1972; Weinreich, 1953). However, a change of perspective was suggested by Kellerman (1995), who identified the role of "cross-linguistic influence" (used interchangeably with "language transfer") as being "considered nowadays as much facilitative as it was formerly thought obstructive" (p. 126).[14] More recently, an even stronger stance has been brought forward by Ringbom (2007), in characterizing transfer as "predominantly positive" and redefining negative transfer as the "*absence of relevant concrete (positive) transfer, leading to subsequent wrong assumptions about cross-linguistic similarities* […]" (p. 30, italics his). However, he also agrees with the frequently

---

[14] Even though the terms 'positive' and 'negative' suggest value judgments concerning good and bad language use, not all instances of mixed language, i.e., the most easily observable kind of (negative) transfer, are necessarily inappropriate.

stated observation that positive transfer is far more difficult to measure which, in turn, explains the scarcity of such investigations in language acquisition research (e.g., Falk & Bardel, 2010, p. 199). Finally, as far as traditions and trends are concerned, a rising interest in TLA and multilingual development over the past fifteen years (see above) has broadened the scope of transfer research with respect to the influence that several non-native languages may have on each other, also referred to as *lateral* transfer (Jarvis & Pavlenko, 2008, p. 20; Odlin, 2003, p. 477). In contrast to L2 learners who only have their L1 to draw upon, L3 learners are able to access more than two linguistic systems and thus potential transfer combinations are "proportionate to the number of languages known to the speaker" (De Angelis & Selinker, 2001, p. 45). However, outcome (positive vs. negative) and directionality of crosslinguistic influence are not the only aspects of language transfer that have been addressed in the literature. Since transfer processes usually don't affect all areas of language, a distinction between phonological, lexical, syntactic, morphological, and, in some cases also discursive, pragmatic and sociolinguistic transfer may be appropriate, according to the scope of the respective investigation (Jarvis & Pavlenko, 2008; see also Odlin, 1989, 2003).[15]

To better situate our study, we will concentrate on lexical transfer, given that the similarities between English and German that could serve as a basis for transfer are most obvious in the lexicon. One distinction that has been quite useful to classify types of lexical transfer is the one between *transfer of form* and *transfer of meaning* (Lindqvist, 2010; Ringbom, 2001). In a later contribution, Ringbom adopted the terms "item transfer" and "procedural transfer" to refer roughly to the same categories (Ringbom, 2007, p. 54). Examples of item transfer could be the use of false friends or coinages, both of which are based on the perceived formal similarities between the source and target language (p. 55). Procedural transfer, on the other hand, refers to semantic or syntactic properties of words transported from one language into the other, such as, for instance, in calques or semantic extensions. An example for the latter type is "He bit himself in the language", where *kieli* is Finnish for tongue and language, respectively (example from Ringbom, 1987). It has been

---

[15] Jarvis & Pavlenko (2008) suggested a ten-dimensional taxonomy to classify language transfer extensively. However, for the purpose of our study, it didn't seem necessary to specify all the possible aspects to be taken into consideration.

shown that learners at initial stages of L2 and L3 learning rely more on formal correspondences between the lexical items they have at their disposal, while more proficient learners tend to make more meaning-based connections between words and other lexical patterns (e.g., Fitzpatrick, 2006; Lindqvist, 2010; Ringbom, 2007). Hence, the type of lexical transfer we will be investigating is expected to be mostly item-based, since our initial-stage learners will have to rely on formal similarities to other languages to make sense of unknown words in their L3.

2.4.2 Prevailing factors affecting (lexical) transfer

In an extensive discussion on methodological rigor in transfer studies, Jarvis (2000) discusses a number of influential factors, whichshould be taken into consideration and, ideally controlled for, in any study of language transfer.[16] Given the scope of our project, we will mostly concentrate on language- and learner-related factors, with a special focus on *typology* and *proficiency,* given that these are clearly the ones to which most attention has been drawn, in contexts of both L2 and L3 learning. Naturally, Jarvis' (2000) categories, when applied to multilingual contexts, necessitate an extension to source languages other than the learner's L1. That is, proficiency in the target but also in the source languages has been shown to affect transfer processes, as will become clear in the subsection related to these factors (2.4.2.2). Subsequently, the usage-based factors of *frequency, recency and exposure* will be presented. Finally, we will briefly discuss some *context-related factors* that may be influential in transfer research. Given the increasing number of studies on lateral transfer and the specific focus of our project, we will put a special focus on studies carried out in multilingual contexts.

2.4.2.1 Language-related factors: Typology and L2 status

As numerous studies have shown (see below), learners of languages that are typologically closer to their mother tongue make greater gains in listening and reading comprehension tasks than those whose mother tongue has little or no resemblance to the target language structure(s). In one of his earlier contributions, Ringbom (1992) speaks of "potential

---

[16] Even though he puts a special focus on L1 influence in L2 or L3 acquisition, the enumerated features are equally applicable to contexts of lateral transfer.

vocabulary" (p. 88) as he discusses the positive influence of cognates for comprehension when reading in an L2. He compared Finnish- and Swedish-speaking adolescents' reading performances in English. The latter outperformed the former significantly which was, according to Ringbom, due to "the Swedes' L1-based potential knowledge [being] more relevant to a learner of English than is that of the Finns" (p. 92). At the same time, not only the L1 but any genetically related language accessible to the learner may provide typological cues which can accelerate the learning process. However, it doesn't only depend on objective similarities and differences. The learner's perception of the proximity or distance between languages has been identified as equally important for transfer phenomena (Cenoz, 2001; De Angelis & Selinker, 2001; De Angelis, 2005; Odlin & Jarvis, 2004; Rothman, 2011; Tremblay, 2006; Williams & Hammarberg, 1998). Perceived similarities and differences have been subsumed under the notion of "psychotypology" (Kellerman, 1977). When looking at transfer phenomena, typology and psychotypology can hardly be separated since it always depends on the particular learner whether similarity-relations between potentially transferrable items are being established or not. Generally, if an unknown word looks similar or identical to a known word in another language, learners are led to believe that there is also equivalence in meaning. However, this may not be the case. Thus, learner assumptions can either be deceptive or greatly advantageous. According to Ringbom (2007), "[w]hen both formal and functional similarities can be established, this makes for positive transfer" (p. 6). And, as stated by Odlin (2003), "[h]aving such opportunities, however, does not guarantee that any particular learner will do the necessary looking or come to the right conclusion about just how congruent a cross-linguistic correspondence is" (p. 443). Hence, typology should not be seen as a free ticket to positive transfer, given that it largely depends on each individual learner whether or not similarities are perceived as such and interpreted accordingly.

Another language-related factor that equally depends on the learner's perception of their language resources is what Williams and Hammarberg (1998) have labeled *L2 status*, which assigns a specific role to the L2 as a source language, especially with respect to lexis. According to their observations of multilinguals using one of their non-native languages, the source language from which transfer is expected should be the one that scores highest

on all four of the following categories: *proficiency*, *typology*, *recency* (see below), and *L2 status*. For example, L2 status might explain transfer from a particular L2, rather than from the learner's L1 (Williams & Hammarberg, 1998, p. 322). Other researchers have supported the claim that learners of third or additional languages tend to resort to their L2 rather than to their native language when using their L3 (De Angelis, 2005; Singleton & Ó Laoire, 2006).[17] According to De Angelis (2005, 2007), two assumptions on the part of the learner could explain this tendency, namely their *association of foreignness* and their *perception of correctness* (De Angelis, 2007, p. 27). On the one hand, if in doubt about a particular TL item, a speaker may opt for another foreign language item instinctively, given that recourse to the native language would clearly be the incorrect choice. On the other hand, the influence from another non-native language is also determined by what the learner perceives as target-like. More specifically, this means that learners may use non-target items or lexical inventions assuming that they are part of the TL system. Based on data from an older study (De Angelis & Selinker, 2001), De Angelis (2005) introduced the term "system shift" to refer to such cases of unconscious cross-talk. She also made an interesting point concerning the identification of the source of transfer in cases of *system shifts*. The example given was that of a native speaker of English who was interviewed in her L3 Italian in which she incorporated numerous lexical items from her L2 Spanish such as *pintura* (Spanish for *paint*) or *nieve* (Spanish for *snow*) instead of the corresponding target items *pittura* and *neve*. Apart from such language switches, she also invented words such as **abbastante*, based on both Italian (*abbastanza*) and Spanish (*bastante*). Interestingly, the subject reproduced the same non-target words in a second production, even though she claimed later on that she was familiar with the corresponding target items in Italian. These and other findings were taken by the researcher to indicate that while performing in a given L3, speakers may take the non-target items they produce to be part of the TL system, when they are clearly influenced by another non-native language (De Angelis, 2005, pp. 6–8). In sum, the reasons that may lead multilinguals to resort to another non-native language while using their L3 seem to stem, at least partly, from their perceptions of correctness of the given TL.

---

[17] It should be stated though, that *L2 status* was also found to affect transfer of syntactic patterns (e.g., Bardel & Falk, 2007).

2.4.2.2 Learner-related factors: Proficiency, frequency/recency of use and exposure

A common assumption is that language transfer, in general, occurs mainly at low levels of TL proficiency. This stance has been supported by researchers who found a decrease of crosslinguistic influence in the learners' use of the respective TL with an increase of proficiency in that language. With a special focus on language switches, Williams and Hammarberg (1998), for instance, made this observation based on oral data of a multilingual learner of Swedish, collected over a period of two years.[18] Accordingly, lexical transfer could be seen as a sign of an incomplete TL lexicon (Wei, 2006, p. 97, see also Cenoz, 2001). However, such a view seems to be an oversimplification. Not only has lexical transfer also been shown to occur at higher levels of TL proficiency (e.g., Bardel & Falk, 2007; Ecke, 2008), but it also seems as though the proficiency factor determined, at least partly, the different kinds of lexical transfer observable in TL use. Ringbom (2007), who provided data on both language production and comprehension (see also studies from 1992, 2001, and his major contribution from 1987), identified a large number of transfer phenomena as ways to make up for gaps in TL knowledge, i.e., as a sign of an incomplete TL system. However, he also stressed the fact that learners at lower levels of TL proficiency tend to transfer lexical items based on formal similarities and differences between the languages. This tendency is explained as a result of stages of vocabulary development, since he argues that language learners at initial stages of TL acquisition tend to rely on formal representations of words, while their lexical knowledge develops into a more semantically-based system as they learn more about the use of each item in different communicative situations (Ringbom, 2007, pp. 54–56). Within the same framework, Lindqvist (2010) investigated lexical transfer in advanced learners of French and argued that crosslinguistic influence was not simply a matter of a more or less developed TL system. She actually found that learners at higher levels of TL proficiency tended towards more semantically-based transfer, while low TL proficiency generally led to more form-based associations between words. Thus, Ringbom's observations concerning a shift from item-based towards procedural transfer have been confirmed in her data. However,

---

[18] However, this tendency wasn't only reported with respect to a decrease of mixed language over time but was also related to phonetic patterns adopted from English at initial stages of learning Swedish.

Lindqvist (2010) included *slips of the tongue*, clearly instances of *unintentional language switches*, in the category of *semantic* transfer. A different categorization would have changed the observed tendency significantly, since it would have indicated a preference for *formal* transfer, even with advanced learners. Nevertheless, studies on word association behavior have also shown that associations based solely on formal properties of lexical items are more frequent with non-native speakers than with native speakers (e.g., Fitzpatrick, 2006), which in turn confirms the above-stated assumption that there may be a shift from formal towards semantic transfer as TL proficiency increases.

Other challenging perspectives concerning TL proficiency have been provided in studies on reverse transfer, i.e., from a second or third language into the learner's L1 (e.g., Ecke, 2008; Jarvis, 2003), since they clearly indicate that lexical transfer doesn't necessarily occur into a low-proficiency language. Finally, we will conclude this first subsection with another observation that should sum up how an increase in TL proficiency is not expected to affect lexical transfer quantitatively but rather qualitatively.

> Becoming more proficient in another language increases the speakers' communicative potential and scope but it also increases the burden put on the speech processor to find appropriate words in a larger lexicon that could be subject to crosslinguistic interaction between language representations and access routes. (Ecke, 2008, pp. 516–517)

As far as the role of source language proficiency is concerned, the same tendency has been reported in a large number of studies on lexical transfer from both native and foreign languages. Generally speaking, high proficiency in a given source language seems to be one of the determining factors for that particular language to interact with the given TL. Thus, speakers of three or more languages tend to resort to a high-proficiency source language, either intentionally to solve communication problems or unintentionally as a result of interference (de Bot, 2004; Falk & Bardel, 2010; Lindqvist, 2010; Odlin & Jarvis, 2004; Odlin, 2003; Peyer et al., 2010; Ringbom, 2007; Tremblay, 2006; Williams & Hammarberg, 1998). However, this doesn't mean that the source language is by default the learner's L1, i.e., the language in which he or she is most proficient. Not only do many influential factors interact in the process of lexical retrieval, but it is possible that the L1

may be more easily inhibited as a whole than non-native languages when learners are communicating in a given L3 (de Bot, 2004, p. 11, see also section on L2 status below). In order to explain problems in TL production, Green (1986) proposed a distinction between a "selected", an "activated", and a "dormant" position that languages can take in the speaker's mind. These different positions depend on regularity in language use and on control mechanisms. Even though one language may be selected for output, the other language system(s), if on an equal or a higher stage of activation, can interfere with the selected system (Green, 1986, pp. 215–216). According to de Bot (2004), there seems to be a threshold of language proficiency for a language to compete in the selection process. Next to other factors that may lead to the co-activation of several languages, the importance of language proficiency has been stated in numerous studies. Dewaele (1998, 2001), for instance, investigated the lexical inventions of Dutch-speaking learners of French, who were also learning English as their L2 or L3, respectively. Both studies revealed that French L3 speakers tended to create lexical items based on their L2 English, while learners who had started to learn English after French, were more inclined to transfer from their native language Dutch. While Dewaele stressed the order of acquisition as the determinant factor for this tendency, the level of proficiency in the given source language seems to play an equally important role. Thus, in line with Williams and Hammarberg's (1998) observations on what they call *L2 status*, one could assume that speakers of more than two languages tend to resort to their L2 rather than to their L1 when communicating in a third or additional language, unless they aren't proficient enough in that other foreign language for it to reach a sufficient level of activation and thus be available for retrieval.

Alongside proficiency levels, *frequency* and *recency* have also been discussed as likely to influence activation levels of elements from a given language to be selected for output. Even though Williams and Hammarberg's (1998) taxonomy has often been cited in L3 contexts, the *recency* factor has, to our knowledge, not often been investigated independently. Exceptions are Dewaele (1998, 2001), Bayona (2009), as well as Odlin and Jarvis (2004) who discussed this aspect with respect to order of acquisition. On the whole, their findings reflect to Green's (1986, 1998) observations regarding the activation of different languages in the learner's mind, for which frequency of use seems to be crucial.

Even though *recency* and *frequency* cannot be regarded as synonyms, both do refer to enhanced levels of language use, which may, in fact, also occur at lower levels of language proficiency, depending on the measure being used. Another interesting distinction between language proficiency and use has been made by Tremblay (2006). She isolated source language *exposure* from *proficiency* and found that the former seemed to have greater impact on the learners' activation of French L2 as a source for lexical transfer by English-speaking learners of German L3 (Tremblay, 2006, pp. 116–117). Thus, only those subjects with *high exposure* to French combined with *high proficiency* produced lexical inventions and metalinguistic comments or questions in French, whereas unintentional switches were the only ones to occur from *low exposure/high proficiency* candidates.[19] These results were taken to indicate that a certain level of automaticity, as reflected in the measure of *exposure*, needs to be attained for an L2 or Lx to play a role in the use of an additional foreign language (see also Segalowitz, 2003, pp. 388–390). Even though Tremblay's findings indicate an interesting perspective with respect to a further specification of what *proficiency* in a given language may also imply, it should be noted that with a total of 13 participants, her results can hardly be taken as general tendency. Furthermore, she pointed out that it might have been more fruitful to conduct the study with French-speaking learners of English L2 and German L3, since the typological similarities between L2 and L3 would have led to more transfer and thus, given way to a more thorough investigation of L2 exposure as a determinant factor, regardless of L2 proficiency.

2.4.2.3 Context-related factors

As far as the learning environment is concerned, few researchers have focused on factors such as the *formality of context*, in which the study was being conducted, or the *work language* that was being used to give instructions. An interesting exception is Dewaele (2001) who investigated the oral French (L2 or L3) production of Dutch university students in a formal and an informal speech situation, respectively. Just like the researcher, all of them spoke Dutch, English and French, yet at different levels of proficiency. As expected from the model of *language mode* proposed by Grosjean (1997, 2001, 2013), mixed language was significantly less frequent in the formal situation. Thus, the students adapted

---

[19] There were no participants with high exposure and low proficiency.

their use of non-target items to the degree of formality of the speech situation. Given the awareness that the researcher was proficient in all three languages, they could have switched freely between the systems, had there not been the constraint of the exam situation where students knew they would be assessed on correctness. Thus, they tried harder to stick to the base language French and only seemed to loosen control and move closer towards the multilingual end of the continuum in the informal situation.

Another context-related factor worth mentioning is the *type of task* that is being used in the investigation of lexical transfer. As pointed out by Jarvis (2000), this might have an enormous impact on the results to be produced. For instance, in a translation task into or out of a given TL, it will depend on the text or the words chosen by the researcher how many possibilities for lexical transfer there are. In turn, language use is manipulated and results must be interpreted with care. The same goes for any other kind of data collection that is not carried out in a naturalistic setting (Jarvis & Pavlenko, 2008, pp. 189–190). Other examples of this kind have been reported for word association tasks, in which the chosen stimulus words greatly influence learners' responses. For instance, high frequency items tend to trigger fairly predictable association patterns, which in turn may not be representative for the whole lexicon. Moreover, researchers rarely use isolated items but rather lists of stimuli and thus have to bear in mind that each item in the list may also influence any subsequent association (Fitzpatrick, 2006, p. 124). Examples like these indicate that lexical transfer may also be affected by the context in which the investigation takes place. Consequently, if factors such as *work language* or *task type* are not measured as influential variables, researchers should account for possible misinterpretations of results due to context-related factors (e.g., Jarvis, 2000, p. 272). As for the specific context in which our study will be conducted, both these aspects have been taken into consideration as a possible bias for the manipulation of the target language material. In a detailed description of the research instruments (see Chapter 4), it will become clear how our methodological choices are motivated with respect to the above observations.

2.4.3 Interim summary: Predicting lexical transfer from L2 to L3

As a common phenomenon of second and foreign language learning, lexical transfer can

take place between all the languages in a learner's mind. It can be either *positive* or *negative*, even though none of these labels should be taken as exclusive, since dimensions of word knowledge are necessarily interrelated. Furthermore, so-called negative transfer doesn't necessarily hinder communication. Out of the factors that have been found to affect lexical transfer, *proficiency* and *typology* have proven to be the most influential. However, especially with respect to L3 learning, other language-related factors such as *L2 status* and *frequency* of use have proven to be crucial for the activation of a given L2 during L3 use. As far as language proficiency is concerned, several hypotheses put forward in the literature have been confirmed. First of all, lexical transfer does not only occur at low levels of TL proficiency, the compensation for a lack of TL vocabulary knowledge with that of previously or simultaneously acquired systems being only one possible explanation for lexical transfer. At the same time, TL proficiency clearly does make a difference with respect to qualitative aspects of lexical transfer. Namely, learners at initial stages of L2 and L3 learning rely more on formal correspondences between the lexical items they have at their disposal, while more proficient learners tend to make more meaning-based connections between words and phrases. Given that formal similarities appear to activate words from different linguistic systems, regardless of semantically or conceptually based subtleties in use, lexical transfer based on lexemic (or formal) dimensions of word knowledge is likely to occur at any level of target language proficiency, while procedural (or meaning-based) transfer seems to be restricted to more advanced proficiency levels in the respective source language. Consequently, source language proficiency plays a crucial role for the activation of that language in the learning process of an additional one, since learners tend to make use of their strongest languages to understand and produce new linguistic material. However, this doesn't mean the L1 is necessarily the primary source of transfer, simply because it's the strongest: rather, when browsing through their lexicon for the appropriate items and structures to choose, learners of two or more languages tend to avoid their L1, which is, at least partly, due to what has been called *L2 status* (e.g., Singleton & Ó Laoire, 2006; Williams & Hammarberg, 1998). Finally, different dimensions of language use naturally play a role for the activation of a given source language. Thus, words from a language that is frequently used will be more easily accessible and thus compete for selection, which may result in lexical transfer (de Bot,

1992, 2004; Green, 1986, 1998). The same tendency was stated for higher levels of *exposure* to a given source language. Interestingly, however, the latter variable has been found to play a role even at lower levels of source language proficiency (e.g., Tremblay, 2006). As far as context-related factors are concerned, they can definitely affect the activation of different languages and thus lead to transfer, however, as the largest body of research suggests, probably to a lesser extent than proficiency, typology and regularity of language use.

In the light of what has been said about the particularities of multiple language learning (see sections 2.2 and 2.3) other variables related to individual multilingualism are similarly expected to affect positive transfer. It should be noted, however, that cognitive abilities of individual learners have clearly not been the focus of study, even though the increasing number of studies on the role of metalinguistic awareness for multiple language acquisition (see above) indicates a promising direction for research on positive effects of lexical transfer in relation to enhanced levels of metalinguistic awareness (Caspari & Rössler, 2008, pp. 71–72; Franceschini, 2009, p. 36; Jarvis & Pavlenko, 2008, p. 196; Jessner, 2006). Namely, an enhanced capacity of analysis and control of multilingual speakers based on their language learning experience (e.g., Bialystok, 2001a) might facilitate the development of L3 vocabulary, since these learners are expected to make more efficient use of their lexical resources (e.g., Jessner, 2006; Singleton & Aronin, 2007; Volgger, 2010). In fact, when considering the factors affecting lexical transfer, a question that clearly deserves further attention is whether metalinguistic awareness may be regarded as an independent variable or whether it is necessarily related to language proficiency. Finally, when it comes to affective variables, such as attitudes towards TL communities, they have, to our knowledge, not been investigated as influential factors in any study on lexical transfer, even though they might be expected to play a role for the accessibility of certain lexical items or systems in a learner's mind (see Hufeisen, 2000).

Taking these aspects into consideration, the present study aims at investigating the roles of source language proficiency, source language exposure and use, and metalinguistic awareness as the main explanatory variables for positive lexical transfer from English (L2)

to German (L3), two West Germanic languages that are typologically closely related and thus share numerous formal similarities, especially on the lexical level.

3   Review of the literature on positive transfer in TLA

Even though multiple language acquisition is commonly believed to have facilitating effects on subsequent language learning due to increased cognitive abilities (see above), researchers investigating crosslinguistic influence have tended to focus on interference-based learner errors such as language switches, lexical inventions or morphological adaptations, even in multilingual contexts (e.g., Cenoz, 2003b; Dewaele, 1998; Ecke, 2015; Lindqvist, 2009; Llama et al., 2010; Tremblay, 2006; Wei, 2006; Williams & Hammarberg, 1998). Only a few studies (see below) have explicitly addressed the positive effects of bi- and multilingualism on the acquisition of a subsequent language. This section aims at reviewing some of these studies in more detail. They will be presented and discussed according to the most prominent factors leading to positive transfer in the particular context of each of the studies. It will thus become clear which language- and learner-specific variables ought to be taken into consideration when investigating positive language transfer in contexts of multiple language acquisition.

3.1   Studies focusing on typological similarities

Several researchers investigated the positive influence of prior foreign language knowledge on the comprehension of a typologically related target language, when that particular target was unknown to the participants. Both Gibson and Hufeisen (2003) as well as Rast (2010) used a translation task from the respective target language that none of the learners were familiar with. In Gibson and Hufeisen's (2003) study, 36 adults some of whom were native speakers of German who learned English as an L3, while others came from different linguistic backgrounds and learned German as an L3, had to translate a Swedish text into their respective L3. Whereas none of the participants had knowledge of Swedish, both English and German are closely related to Swedish. In addition to the written translations produced by the participants, the researchers collected secondary data by means of a post-task questionnaire in order to document the influence that prior language knowledge and context had on the completion of the task. Thus, the primary quantitative data, which were

analyzed for accuracy and error rates, could be related to the participants' comments on helpful source languages and strategies they had used to make sense of the unknown text. The results of the translation task in combination with the bio data revealed that typologically related background languages were activated more easily as a source of positive transfer, which was confirmed by the answers of the post-task questionnaire. Moreover, learners with a larger number of background languages, typologically related or not, produced more accurate translations of the short text in Swedish. The authors interpreted this result in terms of heightened levels of metalinguistic awareness, related to individual language learning experience. This aspect will be discussed in more detail in section 3.3. With regard to the role of typology in positive transfer, Rast (2010) found results similar to Gibson and Hufeisen (2003). In this study, 35 participants translated simple sentences from Polish, which was also entirely unknown to them, into their native language French. As hypothesized, the participants who spoke Russian as an L2 reached significantly higher scores in the translation task than those who did not. These results thus indicated the same tendency: that knowledge of a typologically related L2 or Lx has a facilitating effect on the comprehension of a given foreign language, even when the latter is completely new. In sum, learners tend to rely on prior foreign language knowledge if formal correspondences can be established.

Additional research in which typology was shown to be crucial to transfer processes in L3 learners, is that of Rothman and Cabrelli Amaro (2010) which led to the so-called Typological Primacy Model (Rothman, 2011). In both studies, the focus was on transfer of syntactic structures in initial-stage L3 learners. Whereas the former investigated the null-subject structure of Spanish and Italian as opposed to English or French, the latter was concerned with adjectival placement in Romance languages in contrast to English. What the results of both studies revealed is that the syntactic structures under observation were transferred from the typologically related source language into the respective L3 no matter whether that source language was the learners' L1 or L2. Based on these observations, Rothman concluded that typology, or psychotypology, is the strongest factor to determine L3 syntactic transfer (Rothman, 2011, p. 111). Moreover, he thereby challenges the *L2 status* factor (as discussed under 2.4.2 among the factors affecting transfer), which has also

been supported with respect to syntactic transfer (Bardel & Falk, 2007). In fact, from the observations of Bardel and Falk's (2007) experiment involving the placement of negation in verb-second (V2) languages, such as Swedish, Dutch, and German, it is not entirely clear to what extent typology might have played a more important role than expressed by the authors. They collected data from two groups, who were initial-stage L3 learners of either Swedish or Dutch. In each group, roughly half of the learners had another V2 language with the same syntactic pattern for the placement of negation (either Dutch, Swedish or German) as their L1, while the other half had one of these languages as their second. Even if the results revealed that only the learners with a V2 language as L2 transferred the syntactic pattern under observation, there was another factor related to language background that might have, at least partly, explained this tendency. More specifically, the participants with a V2 language as L1 had English as their second language, whereas the ones with a V2 language as L2 had Italian, Albanian or Hungarian as L1. Even if English does not follow the V2 pattern of the other Germanic languages under observation, it is still typologically related to all of them. Thus, participants may have perceived English as very similar to their respective L3, thus establishing a stronger connection between, say, English (L2) and Dutch (L3), than between Swedish (L1) and Dutch (L3). In other words, the observation that native speakers of a V2 language who learn a very similar V2 language as their L3 did not transfer a particular syntactic pattern from their L1 may be due to psychotypology and not necessarily to the fact that the L2 presumably blocks the access to L1 transfer in L3 acquisition, such as suggested by the L2 status factor (see Bardel & Falk, 2007, p. 480). Such interpretations would have to be confirmed by investigations into the learners' perceptions of similarity between their languages as they are manipulating them, which might further encourage the use of introspective data on transfer processes.

In sum, typology or psychotypology, seems to be one of the major language-inherent factors to explain positive transfer. However, other variables were shown to come into play. In the case of *lateral transfer* (see section 2.4.1 on the notion of language transfer), it appears that the proficiency level in the source language can also be related to "interlingual

identifications"[20] leading learners to resort to their L2 rather than to their L1 in order to make sense of new target language material (e.g., Williams & Hammarberg, 1998).

3.2   Studies focusing on source language proficiency

In line with many of the studies discussed previously, Odlin and Jarvis (2004) investigated positive transfer between typologically related languages, in this case Swedish and English. The study was conducted in Finland, where 6% of the population have Swedish as their first and Finnish as their second language, while the vast majority speak Finnish (L1) and learn Swedish as their L2 before learning English as the habitual L3. In this particular linguistic setting, the aim of the study was to find out whether the influence of Swedish as a source language would affect the acquisition of English differently, depending on the status of Swedish as either L1 or L2. The participants were aged between 11 and 16 and had either Finnish (n = 140) or Swedish (n = 70) as their L1, and were all learning English as their L3. Six experimental groups of 35 were formed, according to age group and respective levels of L2 and L3 proficiency, measured in terms of years of language instruction, which varied between two to six years for the respective foreign languages. One group of the Finnish-speaking learners had not received any Swedish instruction at all. Data were collected by means of a writing task in English, which consisted of a written description of an 8-minute segment of a silent Charlie Chaplin movie. The main focus of the study was on the production of four specific lexical items of English, namely *instead*, *for*, *some* and *what*, all of which have Swedish cognates. Possible mappings of form-meaning correspondences could be explained in terms of language combinations as well as the participants' respective levels of proficiency in Swedish, which was the major source language under observation. The overall results of the study clearly indicate that prior knowledge of Swedish helped learners of English to produce correct written texts in their L3 to different degrees according to whether Swedish was their L1 or L2. Even though positive transfer from Swedish was evident in both cases, learners who spoke Swedish as their L1 mostly used the English words under observation in the appropriate contexts. For instance, in the case of *for*, both Swedish and Finnish speakers with knowledge of Swedish produced the item abundantly, whereas Finnish-speakers who did not know any Swedish

---

[20] This term was first used by Weinreich (1953, p. 7) and later taken up by Selinker (1972, p. 211).

very rarely did, which was taken to reveal positive transfer from Swedish *för*. However, there was a difference in the use of *for* depending on the respective levels of Swedish proficiency, since Finnish-speaking learners of Swedish often overgeneralized uses of *for* according to syntactic patterns of *för* in Swedish, which do not apply in the case of its English cognate. In sum, the results of the study revealed numerous instances of positive transfer from Swedish to the target English. The discussion of the results with respect to correspondences between Swedish and English provided sufficient evidence that the qualitative and quantitative differences in the use of four specific lexical items in English were mainly attributable to the respective levels of Swedish proficiency.

Similar observations were made by Dewaele (2001), who investigated the lexical inventions of Dutch-speaking learners of French, who were also learning English as their L2 or L3, respectively. Both studies revealed that French L3 speakers tended to create lexical items based on their L2 English, while learners who had started to learn English after French, were more inclined to transfer from their native language Dutch. While Dewaele stressed the order of acquisition as the determinant factor for this tendency, this aspect appears to be in close relation to the level of proficiency in the given source language. Moreover, these findings also corroborate previous research such as, for instance, Williams and Hammarberg's (1998), whose qualitative data collected within a longitudinal case study revealed the participant's preference to resort to a high proficiency L2 (in this case German) searching for appropriate words in the L3 (Swedish), rather than to another L2, in which the learner was less proficient.[21] Finally, the above-stated findings reflect de Bot's (2004) observation that there seems to be a threshold of language proficiency in order for a source language to compete in the selection process.

What follows from the above observations is that learners in multilingual contexts, that is "involving the acquisition of any language beyond a first or a second" (Odlin & Jarvis, 2004, pp. 123–124), draw upon their prior foreign language knowledge, especially when the language(s) known to the learner are typologically related and perceived as being

---

[21] In this study, source language proficiency was only one out of four factors (among *typology*, *recency* and *L2 status*) that were shown to affect transfer rates.

similar to the target language. Moreover, instances of positive transfer seem to be particularly frequent, when interlingual identifications are being made on the grounds of a sound knowledge of the given source language. Furthermore, enhanced levels of metalinguistic awareness seem to make learners more efficient in lexical inferencing, as will be highlighted in the following section.

3.3   Studies focusing on metalinguistic awareness

Given that MLA is the key concept of the present study and that it has been approached from different angles and within different research paradigms, this section is separated in two parts. Whereas the first subsection is dedicated to studies having focused on MLA as a predictor of positive transfer in L3 learning contexts, the second involves research where the spotlight is on the development of literacy skills, of which MLA is thought to be an outcome, which, in turn, has a facilitating function on the acquisition process of additional languages. Hence, instead of focusing on MLA as a driving force in itself, the studies presented under 3.3.2 rather target its role as a mediator between the mental processes related to bilingualism and SLA and the learning of subsequent languages.

3.3.1   MLA as a predictor of positive transfer

The studies on positive transfer in which MLA was directly addressed as an influential factor have mostly approached this aspect qualitatively by using intro- or retrospective methods of data collection. To our knowledge, Jessner's (1999) was the first introspective study to investigate the role of MLA for the use of crosslinguistic cues in an L3 production task. Embedded in the above-stated dynamic model of multilingualism (Herdina & Jessner, 2002), she underlined the crucial role of MLA for an accelerated learning process of additional foreign languages. In order to stress the (presumed) cognitive advantages of multilingual learners over their monolingual peers, she evoked Cook's (1992) notion of multicompetence and also mentioned previous studies in bilingualism research such as, for instance, Nayak et al. (1990), in which learners/speakers of more than one language were shown to be more flexible regarding their use of language-learning strategies than monolinguals. Setting the stage for her discussion from this perspective already suggests that MLA is viewed as a basis for strategy use in an L3 learning context.

Jessner's methodology consisted of a written text production task in English (L3), during which the participants, namely Italian-German bilingual university students in Austria[22], were asked to produce think-aloud protocols. These introspective data were used to unveil the strategies that were being used by the participants to make interlingual connections while searching for words in their L3. The results revealed in which way these experienced learners made use of all their languages in order to produce L3 texts. It thus became obvious that there was a conscious search for similarities between Italian, German and English when participants were in doubt about specific features of the target language. As stated by the author, these data "provide evidence of the metalinguistic thinking involving the usage of all three, typologically closely related, languages in the process of academic writing" (p. 205). In this holistic view of the individual processing of multilinguals, MLA could clearly be identified as the chief factor for positive transfer, which corroborates the claims concerning a positive relation between multiple language learning and metalinguistic development.

Comparable yet less detailed observations have been made by using retrospective methods such as questionnaires used to collect qualitative data on processing mechanisms, namely by Gibson and Hufeisen (2003, see section 3.1). On the one hand, MLA was manifested implicitly via the degree of correctness of the task results, i.e., a translation task from an unknown foreign language (Swedish) into a related target (English or German). On the other hand, the post-task questionnaires were used to elicit the participants' comments about specific strategies to infer meaning such as the recognition of cognates or syntactic structures similar to other languages they were familiar with.

Similarly, Dillon (2009) found evidence for MLA based on self-ratings from trilingual pupils in Ireland who indicated to what extent they noticed crosslinguistic similarities and used their prior language knowledge to learn their L3. In this specific case, the frequency of "associations between L1, L2 and L3" (Dillon, 2009, p. 194) as measured on a five-point scale (*always, often, sometimes, seldom* or *never*) was used to indicate the respective levels

---

[22] The number of participants is not specified in the article.

of MLA. Dillon compared balanced bilinguals (English-Irish) from immersion schools to pupils from English-medium schools, where Irish is taught as an L2 after English. All 160 participants were currently learning French and German as additional languages, and some also Spanish and Italian. The results revealed that balanced bilinguals established associations between their languages more easily and also more accurately, thus indicating that a richer language background might favor metalinguistic development, which was also shown to relate to positive transfer. However, as discussed in the previous section, the tendency of balanced bilinguals to resort to positive transfer more often and more efficiently may also be related to higher proficiency levels in both of their potential source languages, English and Irish. Furthermore, instruction-based literacy practices in different languages were shown to foster metalinguistic abilities, which seem to be transferable across languages (e.g., Durgunoğlu, 2002). This aspect will be discussed in more detail in the following subsection. However, before turning towards the relationship between literacy development and MLA, we will briefly discuss another qualitative study, in which MLA could be related to the felicitous use of related vocabulary in a background language when the participants were faced with novel words in a target language.

Dressler and her colleagues (Dressler, Carlo, Snow, August, & White, 2011) found that MLA in combination with the "ability to articulate a strategy" to infer word meaning was associated with higher accuracy (p. 253). The participants, twelve fifth-grade students in Santa Cruz, four of whom were English monolinguals while eight were Spanish-English bilinguals, were asked to infer the meaning of six low frequency words in English that were of Romance origin and thus typologically related to Spanish, but all unknown to the participants. Half of each group had received explicit instruction of strategies such as the use of context clues, the analysis of word parts and the use of cognates. The analysis of think-aloud protocols during a short reading assignment including the target items revealed that the cognate strategy (CS) was the most efficient to infer word meaning, but that it seemed to be primarily the explicit strategy instruction that led to correct lexical inferencing. More specifically, out of 12 correct answers from Spanish-English bilinguals using the CS, only one answer came from a participant who had not received that specific instruction. Even if the authors do not provide an exact definition of MLA, they speak of an

interplay between declarative and procedural knowledge such as evidenced in their verbalizations (Dressler et al., 2011, pp. 252–253). In other words, not only the explicit knowledge of linguistic facts and rules but also the application of such knowledge, indicated by the use of different strategies, was included in the notion of MLA. Despite the small sample size, these observations further encourage a research perspective that takes into consideration the role of MLA in the active use of background languages leading to positive transfer. Moreover, they point to the fact that explicit instruction in which the learners' attention is drawn to certain linguistic rules that can be applied in other learning contexts can be expected to benefit the learners' effective use of inferencing strategies, such as the use of cognates.

In sum, the studies above indicate that L3 learners make use of their other languages when working on different receptive and productive tasks in their L3, especially when the given background languages are typologically close to the target language. Moreover, introspective methods such as think-aloud protocols or retrospective methods such as questionnaires or interviews, appear to provide rich data on metalinguistic processing, and shed light on the importance of individual experiences related to language learning and how these may impact the kind of links established between languages. Even if such instruments do not allow for a systematic measurement of MLA in terms of a quantifiable predictor, these qualitative data provide evidence for crosslinguistic consultation as a learning strategy, which clearly reflects metalinguistic abilities. In the next subsection, the role of MLA in L3 learning contexts will be discussed with respect to literacy development. In particular, while MLA may be viewed as a predictor of positive transfer, it has also been treated in terms of a side effect of literacy practices of experienced learners. A closer look at some of these studies will help pinpoint the specific abilities that are expected to enable learners to develop MLA, which might trigger or enhance positive transfer effects.

3.3.2   MLA as a function of literacy development

Much research on multilingual development has focused on learners who are also involved in literacy-based instruction in more than one language (e.g., Burton, 2013; Nilsson, 2009, among many others). In other words, the kind of multilingualism under observation in

psycholinguistic approaches is often also multi-literate. Even if the importance of literacy practices for metalinguistic development has been hinted at in the conceptual framework, a more thorough review of the relevant literature pertaining to this relationship will help specify the role that MLA has been assigned in L3 learning.

In a recent study on L3 reading proficiency, Rauch and her colleagues (Rauch et al., 2012) pointed to some contrasting findings on the positive effects of bilingualism on L3 learning, which had already been discussed in some detail by Cenoz (2003a, pp. 75–78). In particular, whereas investigations into groups of bilinguals who were also literate in both their languages (e.g., Bild & Swain, 1989; Cenoz & Valencia, 1994) revealed such positive effects, other researchers could not confirm this trend for immigrant populations whose first language had never been used in instructional contexts (e.g., Sanders & Meijers, 1995; Schoonen et al., 2002). These observations led Rauch and her colleagues (2012) to the assumption that a lack of literacy development in L1 impacts literacy development in L3. In other words, they claimed that bilingualism is only positively related to L3 reading proficiency if the speakers under observation are also biliterate. Moreover, the rationale behind their study was based on research findings having identified MLA as a key variable in the bilinguals' advantage over monolinguals to learn subsequent languages (Rauch et al., 2012, p. 403). With reference to Cenoz (2003a) who was one of the first to present MLA as a mediator[23] between bilingualism and L3 development (p. 82), Rauch and her colleagues further investigated if the predicted positive effect of biliteracy on L3 reading was mediated by MLA. They looked at 299 grade-nine students from 14 different schools in Hamburg, 158 of whom were German monolinguals, while 141 were Turkish-German bilinguals having grown up in Germany. Out of the bilinguals, those who scored at least B1 (CEFR scale) on reading assignments in both Turkish and German were considered biliterate. All participants were learners of English as L2 or L3, respectively. MLA was measured by means of one subsection of the so-called Language Awareness Test (Fehling, 2008), which required participants to extract morphological and syntactic rules from examples in English,

---

[23] In the conceptual framework, this mediating function of MLA was described in terms of a "reciprocal effect" (see section 2.3.2 on the relationship between MLA and proficiency in additional languages) that is triggered by language learning, which increases levels of MLA, which, in turn, is thought to facilitate further language learning.

Swedish, and Dutch and to apply those rules to Swedish and Dutch, both unknown to the participants.[24] In line with their assumptions, Rauch et al. (2012) found that full biliterates reached significantly higher scores on English reading proficiency than partial biliterates and monolinguals and that this pattern was comparable for their performance on the measure of MLA. Hierarchical multiple regressions revealed that biliteracy predicted L3 reading proficiency as well as MLA, which, in turn, predicted L3 reading. Moreover, the hypothesis of MLA acting as a mediator was confirmed in so far as the positive correlation between biliteracy and L3 reading was reduced when entering MLA into the equation (Rauch et al., 2012, p. 412). Finally, they proceeded to a direct test of mediation to corroborate their results. This led the authors to the conclusion that positive effects of biliteracy on L3 reading are at least partly due to heightened levels of MLA, which seem to be triggered by literacy practices in L1 and L2.

In fact, the above observations corroborate previous studies where it was shown that so-called "meaning-making strategies" such as "rereading, visualizing and invoking prior knowledge", if acquired in one language, will equally be exploited for additional languages and thus facilitate reading comprehension (Durgunoğlu, 1997, p. 262). There are several other studies where the focus has been on the transfer of processing skills in L3 acquisition. This change of focus is particularly interesting in the case of learners whose languages are not typologically related such as, for instance, in a study conducted in Israel which involved speakers of Russian (L1), Hebrew (L2) and English (L3), as well as monolingual speakers of Hebrew (L1) learning English (L2) (Schwartz, Geva, Share, & Leikin, 2007). Schwartz and her colleagues were interested in the transfer of literacy skills such as phonological decoding, word identification and the acquisition of alphabetic concept in languages with different orthographies, in this case Cyrillic, Semitic and Latin. Out of the 68 Russian-Hebrew bilinguals, a total of 50 students had received literacy instruction in specific community classes and were considered biliterate, whereas none of the other 18 bilinguals,

---

[24] Rauch et al. (2012) justified the use of this particular test, which only focused on morpho-syntax, based on Bialystok's (2001b) observations that the level of bilingualism makes a difference for tasks that call for analysis of representational structures (e.g., syntax correction) but on the other hand, that tasks calling for control of attention (e.g., ignoring semantic errors, sun-moon problem, etc.), even bilinguals who are not highly proficient in both languages outperform their monolingual peers (see p. 405).

for whom Russian was also the dominant language at home, "were able to name more than 4–5 letters of the Russian alphabet or read even the simplest Russian words" (p. 32). The monolingual speakers of Hebrew were, of course, also monoliterate. All the participants were tested on different literacy measures in English such as word identification, word attack, identification and spelling of high frequency words as well as pseudo-word spelling at the end of grade 5, after two years of English instruction. Moreover, they were asked to complete a number of linguistic and metalinguistic tests in English, including initial and final consonant isolation, phoneme separation and grammatical knowledge. Besides, different measures of literacy as well as linguistic (mostly vocabulary knowledge) and metalinguistic (mostly metaphonological) skills were also administered in Hebrew and in Russian. The results that are most relevant to the focus of our study are the following: First, the biliterate bilinguals reached higher scores on the Russian metalinguistic measures such as phoneme deletion and naming of words derived from the same morphological root than their monoliterate bilingual peers. Moreover, biliterate bilinguals outperformed both comparison groups on phonemic manipulation, pseudo word decoding and morphological awareness in Hebrew. Finally, there were no differences between groups in consonant isolation measures in English, but biliterate bilinguals outperformed both other groups on phoneme analysis and pseudo word as well as consonant cluster spelling in English. In sum, these results suggest that biliteracy has a positive impact on phonological processing skills in a third language, even if the latter is not typologically related to the given source language(s). Biliteracy also predicted phoneme analysis both in English and Hebrew. Furthermore, biliterates outperformed their monoliterate counterparts on phonemic manipulation in Russian. The superiority of biliterates on L2 and L3 phoneme analysis and decoding skills can be related to these basic literacy skills in their L1. This was also confirmed by a multiple regression analysis where bilingualism alone could not predict English reading and spelling skills whereas biliteracy could (see Schwartz et al., 2007, pp. 43–45).

In other words, literacy skills in a strong source language tend to be a predictor of success for foreign language learning, given that a number of these skills are "known to transfer across languages" (Durgunoğlu, 2002, p. 201). By extension, highly literate L1 readers are

likely to acquire reading proficiency in an additional language more easily than those for whom reading and other literacy practices do not count as regular mental activities. For instance, Peyer and her colleagues (2010), were interested in how the overall reading proficiency of multilingual learners of German having French or Italian as L1 and English as L2 would affect their L3 reading comprehension. A correlational research design was used to investigate "the advantages in understanding and decoding German sentence structure when reading German as an L3" (see title). The authors designed a German specific reading test for 506 French and Italian university students learning German as a foreign language in which they included a number of syntactic and grammatical structures that had been shown to cause the most difficulties in understanding. Furthermore, the study aimed at showing a positive correlation between the students' reading competence in English as well as their "overall-foreign-language reading competence", excluding German (Peyer et al., 2010, p. 225), and their results on the German-reading test. The participants were first tested in German by means of the online German placement test of the Goethe-Institute. Levels of reading competence in English, on the one hand, and in foreign languages in general, on the other, were established on the basis of a self-assessment questionnaire. The overall test results revealed a small but significant correlation between English reading competence and the German reading test. Moreover, after a separate analysis by levels of German proficiency, as represented on the CEFR scale, it was shown that this correlation was only significant for the "advanced beginners" (p. 236). The same tendency was shown for foreign language competence in general. There was a significant difference of understanding between those who considered themselves competent in foreign language reading and those who did not. Yet, this distinction was particularly obvious at a German proficiency level of A2, namely the advanced beginners. As an explanation for this tendency, the researchers suggest for this particular level that "readers possess a basic knowledge of German but still need to fall back on other resources in order to understand complex German sentences" (p. 236). Interestingly, the typological similarities between English and German did not seem to be particularly advantageous for this German reading task. Obviously, German sentence structure differs significantly from that of English, which may be one possible reason why English reading proficiency alone did not appear to be exceptionally helpful in a German assignment involved mainly with syntactic

processing. Rather, it appears to be the general reading skills, both in L1 and in L2, that facilitate the understanding of new structures when reading in an additional language, especially for non-advanced learners of the target language. Finally, the authors refer to Herdina and Jessner's (2002) notion of the M-factor (as discussed under 2.1.2.2 on the Dynamic Model of Multilingualism) in stating that "multilinguals are more used to analyzing various grammatical input which may lead to greater flexibility and an enhanced metalinguistic awareness when reading German as L3" (Peyer et al., 2010, p. 236).

Cenoz and Gorter (2011) adopt a similar standpoint for instructional writing practices based on their exploratory data from the written assignments of 165 trilinguals from the Basque country. Participants were asked to produce short texts of any type based on a different picture in each of their languages: Spanish (L1 or L2), Basque (L1 or L2) and English (L3). First, the composition scores, which had been divided into five categories (content, organization, grammar, vocabulary, and mechanics), were shown be highly correlated between the three language pairs for all categories except 'organization', thus indicating that participants had comparable writing skills in their different languages, even when they were typologically as distant as in the case of Basque in relation to the other two Indo-European languages (see p. 362). Moreover, general writing strategies[25] tended to be the same across languages (pp. 363-365). Finally, qualitative analyses revealed that transfer took place between all the languages and in all directions, thus indicating soft boundaries between the languages of multilinguals.

Based on these observations of multilingual usage in an instructional setting, Cenoz and Gorter (2011) deplore "the traditional strategy of separating languages and using only the target language in a class", which, according to them, "does not allow multilinguals to use a powerful resource in communication" (p. 367). In turn, it is their conviction that "an integrated curriculum can offer the best conditions to develop metalinguistic awareness and enhance the acquisition of different languages" (p. 359). Again, crosslinguistic influence is presented as predominantly positive, especially in relation with literacy practices, which

---

[25] What the authors identified as different "strategies" were, for instance, preferences for a certain text type (e.g., description vs. story) or for specific content elements (e.g., colors, locations, etc.).

were shown to have facilitative effects across languages. And MLA is thought to be the driving force enabling speakers of multiple languages to use their resources in a variety of contexts.

3.4   Interim summary: Investigating MLA as a predictor of positive transfer

From the above discussion it follows that the most basic explanations of positive transfer are related to the objective and perceived similarities between the languages involved. In this respect, typology has proven to be a promising source for positive language transfer insofar as there are chances for language learners to notice such similarities and make use of them in order to produce or understand new items or structures of a given target language. Another important factor to account for positive transfer in TLA is the level of source language proficiency. That is to say, at higher levels of L2 proficiency, learners of an L3 are more likely to use their L2 as a resource for L3 production or comprehension. Finally, the central factor to be investigated as a predictor of positive transfer in the present study is MLA. In fact, apart from mapping linguistic knowledge from one system to another, learners of multiple languages are likely to draw upon general literacy-based skills when engaging in L3 learning practices. In sum, learners who have received literacy-based instruction in more than one language were shown to develop metalinguistic abilities, which they tend to apply to other learning contexts (see Durgunoğlu, 2002; Rauch et al., 2012; Schwartz et al., 2007). In turn, when investigating the role of MLA in terms of a facilitator of L3 learning, it should be born in mind that the active use of metalinguistic abilities is most probable in instructional settings where learners are engaged in tasks that involve literacy skills such as orthographical, phonological or syntactic decoding and where they are not restricted to use the target language only (see Cenoz & Gorter, 2011).

Even if some researchers have stressed methodological difficulties related to the measure of MLA (e.g., Herdina & Jessner, 2002; Singleton & Aronin, 2007), several research designs have been developed to identify degrees of MLA in speakers of more than one language. To our knowledge, the most comprehensive test battery designed for this goal is the Metalinguistic Ability Test (MAT)[26] (Pinto et al., 1999), which has been translated from

---

[26] A detailed description of this test follows in the Methodology section.

Italian into English, Spanish, and recently into French and which has been used in subsequent studies on multilingual learning (e.g., El Euch, 2010; Lasagabaster, 2001; Pinto et al., 2002), yet never with a focus on positive transfer. The goal of the present study is to combine different methods of data collection in order to investigate more thoroughly whether higher levels of metalinguistic awareness, as measured by Pinto et al. (1999), lead students to make conscious use of their L2 vocabulary in order to make sense of unknown words in their L3, as reflected by think-aloud protocols during a translation task. Furthermore, we will address the question whether levels of interest in the target language and culture might affect the conscious search for crosslinguistic cues.

3.5   Research questions and hypotheses

The goal of the present study is to determine which of the variables that have been shown to affect crosslinguistic interaction in multilingual contexts appears to be the strongest predictor of positive lexical transfer from L2 to L3. Whereas previous research has established the highly influential roles of certain language and learner-related variables such as typology and proficiency, that of MLA has, to our knowledge, not been investigated as a predictor variable of positive transfer. It has been argued that multilingual speakers whose proficiency levels differ across languages exhibit relatively stable levels of MLA independently of the language based on which these abilities were assessed (e.g., El Euch, 2010). Moreover, it was shown that literacy skills (a major constituent of MLA) were transferable from one language-learning context to another (e.g., Durgunoğlu, 2002), which further supports the claim that MLA can be treated independently from language proficiency. The challenge of the present study is to answer the following research questions, while isolating the variables of MLA and English language proficiency, in order to indicate their specific roles for positive transfer in this particular context of multiple language learning.

(RQ1) Do French-speaking beginning learners of German (L3) make use of their prior knowledge of English (L2) to discover the meaning of new lexical items in German (L3), thus revealing positive lexical transfer from English?

Based on a large body of research on both positive and negative language transfer, typological proximity plays a major role for crosslinguistic interaction as a part of the learning process (e.g., Ringbom, 2007; Rothman, 2011). Hence, we hypothesize that the French-speaking Quebeckers taking part in our study draw upon their English (L2) vocabulary knowledge in order to understand new words in German, especially when the target items display formal similarities with English.

(RQ2) Does metalinguistic awareness relate to positive transfer from English?

As suggested by a growing body of research in the field of third language acquisition, MLA plays a facilitative role in the learning process, in that it enables learners to efficiently manipulate linguistic resources (e.g., Cenoz & Gorter, 2011). In the specific context of this study, MLA is expected to help learners decode unknown L3 words based on related structures from other languages, especially the typologically related L2, English. More specifically, we hypothesize that MLA predicts positive transfer from English. Moreover, given that a central role has been attributed to MLA in a variety of research contexts in which multilingual learners were shown to make successful use of their various linguistic resources (e.g., Cenoz, 2003a; Forsyth, 2014; Peyer et al., 2010), MLA is expected to be a stronger predictor of positive transfer than L2 proficiency and exposure, respectively.

(RQ3) Does English language proficiency relate to positive transfer from English?

Transfer research has shown that higher levels of source language proficiency generally predict the extent to which a particular language is activated as a source of transfer, both in production and comprehension of target language structures (e.g., Odlin & Jarvis, 2004). Based on these findings, we hypothesize that English language proficiency predicts positive transfer from English.

(RQ4) Does exposure to English relate to positive transfer from English?

Given that individual variables related to language learning experience appear to affect the L3 learning process (see Hufeisen, 2000), and that some studies suggest that the amount of exposure to a given source language may predict transfer independently of source language proficiency (Tremblay, 2006) this variable was examined separately. In line with these assumptions, we hypothesize that exposure to English correlates positively with transfer rates.

(RQ5) Do levels of interest in the German language and culture, respectively, relate to positive transfer from English?

Finally, with respect to the affective dimension of second and third language acquisition, it has been posited that motivational and attitudinal variables can influence the extent to which learners activate their linguistic resources in different learning contexts (e.g., Hufeisen & Gibson, 2003). Based on this assumption, we hypothesize that levels of interest in the German language and culture, which may be a motivating factor for learners to uncover the meaning of new target language items, also relates to positive transfer from English.

CHAPTER 4

4 Methodology

This chapter is devoted to the methodology adopted in the present study. Having outlined the rationale behind the overall approach (4.1), we will provide a detailed description of the participants (4.2). The subsequent subsection includes descriptions and justifications for the instruments used as methods of data collection (4.3). Finally, we will describe data collection procedures (4.4), the data analysis (4.5) and statistical analyses (4.6).

4.1 Overall approach

The present study adopts a dual approach to the investigation of language transfer by combining quantitative and qualitative methods to collect data on crosslinguistic interaction. What quantitative methods have revealed since the early investigations in language transfer, is that learners with typologically related background languages usually outperform learners with non-related background languages in a variety of learning tasks in the target language, ranging from vocabulary questions (Ard & Homburg, 1983) to reading comprehension tasks (Ringbom, 1987, 1992, 2007). In a more recent study (Odlin & Jarvis, 2004), learners had to perform a writing task in L3 and a fine-grained analysis of their compositions revealed how the appropriate use of certain target structures could be traced back to high proficiency levels in a related background language. In sum, inferential statistics may allow researchers to trace back the felicitous use of a target structure to similarities between the learners' source and target languages. Even though the above-mentioned studies mostly investigated advanced learners, the same tendency could be expected with beginners:

> Establishing cross-linguistic similarity relations is particularly relevant for the comprehension of a new language. When both formal and functional similarities can be established, this makes for positive transfer. (Ringbom, 2007, p. 6)

However, as discussed by Kellerman (1995), for example, a closer look at early-stage learners who have related background languages in their linguistic repertoire, has revealed

that they do not necessarily transfer overtly similar structures into their developing interlanguage. More precisely, he refers to qualitative data from different case studies (Giacobbe, 1992; Giacobbe & Cammarota, 1986; Wode, 1977; as cited in Kellerman, 1995, pp. 131-133) where the learners show a high degree of variability in the integration of related source language structures in their target language (TL) performance, even when both formal and functional similarities could have been established. Not only must these potentially helpful correspondences be perceived as such by the individual, but other variables may also affect the use of prior language knowledge in the learning process of an additional language. Several studies in TLA have investigated the conscious search for crosslinguistic similarities from a qualitative approach. The semi-directed interviews with multilingual immigrants in Germany, conducted by Volgger (2010), for instance, revealed that the learners' perception of their own proficiency levels in their background languages affected the degree to which they consciously resort to their L1 or another L2, even when they believed that there were structural overlaps between, say Turkish or German, and the new TL French. In this study, however, there were no linguistic data to show what kinds of transfer these learners actually produce in their French classes. The participants in Jessner's (1999) study (see section 3.3.1), on the other hand, had to produce written assignments in their L3, which were accompanied by think-aloud protocols (TAPs). Hence, what learners said they were doing and thinking could be related to the linguistic material, as reflected in the TL production data. The analysis of the TAPs revealed the participants were systematically resorting to their other languages as they were searching for appropriate structures in the TL.[27] Even if Jessner's study did not directly investigate the presence or absence of positive transfer, the use of TAPs was an effective way of documenting crosslinguistic word searching by focusing on the learners' individual associations, which could not have been deduced from the linguistic data alone. In line with this approach, Gibson and Hufeisen (2003) combined a linguistic task with qualitative data to investigate crosslinguistic interaction in multilinguals. Learners with various language backgrounds had to translate a short text from an unknown language (Swedish) into their respective L3 (German or English), both of which are closely related to the TL. The task was followed by

---

[27] Note that these individuals were German-Italian bilinguals, i.e., highly proficient in both background languages, and advanced learners of English (L3), which partly explains why all participants made interlingual identifications without much hesitation.

a questionnaire in which participants were asked to explain how they arrived at their translations. By analyzing these answers qualitatively, the researchers were able to identify whether correct translations were based on cues from a given background language, thus revealing positive transfer, or whether they were logically related to structural elements of the co-text, or simply a good guess.

Based on the literature that revealed the ways in which learners resort to typologically related background languages to analyze target words and structures in a new language, our main study consisted of a translation task including a number of lexical items in German (L3), which were, for the most part[28], unknown to the learners, while having formal similarities with their L2, English. Correct translations were thus expected to reveal positive transfer from English. To confirm this assumption, we were inspired by Gibson and Hufeisen (2003) to collect secondary qualitative data in order to highlight the individual processes the learners would engage in during the task. Instead of resorting to a post-task questionnaire, we used TAPs (see Jessner, 1999) to serve this goal since concurrent introspection was shown to be more reliable than retrospective probing, given that the information to be retrieved from long-term memory after task completion is likely to be incomplete and/or subject to interpretation (Bernardini, 2001, p. 243). Finally, transfer rates were established on the basis of correct translations, unless participants indicated prior knowledge of the German test item, and their translation could thus not be traced back to influence from English.

4.2   Participants

The participants were francophone Quebeckers between the ages of 17 and 24, recruited from five different *cégeps* across the province of Quebec. At first, we intended to work with eight (8) intact classes with an average of 22 students from the following institutions: three groups from *Cégep Limoilou* (Quebec City), two from *Cégep de Sainte-Foy* (Quebec City) and *Cégep Ahuntsic* (Montreal), and one from *Cégep Beauce-Appalaches* (St-Georges). The participants were first solicited in early September, between their 3rd and 4th

---

[28] Even if the German teachers confirmed that the items in question had not yet been targeted in class, most of them do appear in the different manuals that each of the teachers use in their classes. As a consequence, some participants already knew their meaning.

week of class. After an oral presentation on the nature of the study, its goals and procedures, given by the researcher in each of the groups, the students who were interested in participating were given consent forms, which they were asked to read and sign by the following week. A total of 110 students accepted to participate in the project by signing the consent forms, which were collected the following week. We only retained students who had French as their first and English as their second language, and who were enrolled in an arts program with a language profile (*Art et Lettres – Profil Langues*) in which they had recently started to learn German as an L3, along with another foreign language, Spanish[29], after their L2, English. Most of the participants had started to learn English at the age of eight, i.e., in third year of primary school. With an average age of 18 in each group of students, the participants had received an average of ten years of English instruction (see Table 1 below).

The first analysis of the background questionnaire revealed that a surprisingly large number of students came from different language backgrounds than those targeted in our protocol. Especially at *Ahuntsic College* in the Montreal area, the majority of the participants who had already accepted and started to participate in the study had to be excluded from the sample, since most of them had languages other than French as their mother tongue. Moreover, we had to exclude participants who had grown up with another language besides French or who had had extensive exposure to other languages during their childhood. Participants who had had contact with German in either formal or informal settings before starting the program were also excluded from the sample. Finally, students who had taken German as optional courses to complement a different program of study were excluded from the sample, given that these participants were not necessarily learning German and Spanish simultaneously, nor would they have received as many hours of English instruction per week[30]. The overall aim of this procedure was to ensure high intra-group homogeneity so as to be able to examine the individual factors of proficiency in and

---

[29] Some of the students entering *cégep* had already taken Spanish classes in secondary school, in which case German would have to be considered their 4th language. However, we used the term L3 for all languages beyond the second that were currently being learned and observed.

[30] For instance, at *Cégep Beauce-Appalaches*, students enrolled in *Art et Lettres* with a language profile have to take seven English courses (3 h/week) throughout the 2-year program, as opposed to other programs where only two English courses are mandatory.

exposure to English as well as their levels of metalinguistic awareness (MLA), more exclusively. The remaining number of participants after this filtering procedure and several dropouts throughout the entire period of the study was fifty-eight (58). However, since we had initially aimed for a total of eighty (80) participants[31], we decided to recruit more students from *Cégep de Lévis-Lauzon* at the end of October. Eight (8) additional participants were therefore added to our sample.

Table 1. *Participants*

School	*City*	*Number*	*Male*	*Female*	*Age*	*Years ESL*
Cégep Limoilou	Québec	22	5	17	17-24	9-13
Cégep de Sainte-Foy	Québec	11	1	10	17-20	6-13
Ahuntsic College	Montréal	5	2	3	17-22	9-14
Cégep Beauce-Appalaches	St-Georges	20	2	18	17-21	8-13
Cégep de Lévis-Lauzon	Lévis	8	-	8	17-19	9-13
Total (or mean*)		**66**	**10**	**56**	**18***	**9,87***

At the time of data collection, the participants, except for those from *Lévis-Lauzon,* had received between 20 and 24 hours of formal instruction of German. Their vocabulary knowledge was thus restricted to numbers, personal pronouns, some basic verbs and nouns, and a restricted number of collocations. Even if German vocabulary knowledge was thus expected to be equally limited for all participants, we could not control their degree of personal investment into their learning process during the first few weeks of exposure to this new language. In turn, some students had already come across a certain number of the lexical items that were targeted in our translation task. As indicated in the coding procedures (see section 4.5.1), correct translations that participants explicitly indicated as being based on prior knowledge were not included in the transfer measure. For each participant, codings were normalized accordingly, that is, the total number of instances

---

[31] Prior to data collection and analysis, a statistician suggested a minimal number of 80 participants to yield a valid statistical analysis of our research questions, if we were to run linear multiple regressions. However, once the data set was analyzed and the results presented to and discussed with the statistician, it turned out that the particular pattern of our translation task called for binary logistic regressions. In turn, it would not have been necessary to attain a specific number of observations for this specific kind of statistical analysis.

coded as transfer was calculated in relation to the total number of trials that did not include instances of prior knowledge.

4.3   Methods of data collection

To begin, demo-linguistic data were collected by means of a *Background questionnaire* (4.3.1) including the identification of levels of exposure to English outside the classroom, as well as respective levels of interest in the non-native languages and cultures. Positive lexical transfer was measured by means of a *Word translation task* (4.3.2), which was complemented by *Think-aloud protocols* (4.3.3). In addition, a *Metalinguistic ability test* (4.3.4) and an *English proficiency* test (4.3.5) were used to determine the participants' levels of metalinguistic awareness and of English language proficiency, respectively. Both these variables were analyzed as predictors of positive lexical transfer from English to German. The following subsections provide detailed descriptions of the research instruments in the above-stated order. In order to allow for a systematic presentation of these methodological choices, further subdivisions are devoted to (1) the rationale, (2) the description and (3) the validity and reliability of each instrument. Only in the case of the word translation task, the last subdivision provides additional information on the choice of the test items.

4.3.1   Background questionnaire

4.3.1.1 Rationale for the background questionnaire

The background questionnaire designed for the present study was largely inspired by the "Language Experience and Proficiency Questionnaire (LEAP-Q)" designed by Marian and her colleagues (2007) to assess language profiles of multilingual learners. Following the authors' recommendations, other researchers such as Bayona (2009) have adapted the questionnaire to their specific research context. Given the relative similarity between Bayona's (2009) and our own study regarding the focus on transfer phenomena from L2 to L3 in an exclusively multilingual group, we adopted her version of the LEAP-Q with slight modifications of content, especially regarding different aspects of L2 and L3 exposure, and translated it into French. In the present protocol, the questionnaire was used to sketch out the participants' linguistic profiles with respect to the following aspects:

1. Number and order of previously acquired languages
2. Age and context of acquisition of each of the non-native languages
3. Time spent in four linguistic environments (country/province, family, school, work place) where these non-native languages were spoken
4. Current exposure to their non-native languages in different dimensions of life, related to (mostly informal) language use in oral and written mode
5. Levels of interest in their non-native languages.

4.3.1.2 Description of the background questionnaire

Our background questionnaire consisted of four sections the first of which addressed the order of acquisition of all of the participants' languages, according to their perceived dominance, while the second, third and fourth sections focused on their knowledge and learning experience of English, Spanish and German, respectively. Even though Spanish did not play a specific role in our protocol, insofar as it was not analyzed as a potential source of transfer in the translation task, the data relative to individual learning experience of Spanish were collected in view of subsequent analyses.[32]

Preceding the actual questions, the heading of the questionnaire included a table in which participants had to indicate their name, age, gender, the school they attended and the program in which they were currently enrolled. The first section included five (5) questions, four of which pertained to the order of acquisition of the respective languages targeted in our study: French, English, Spanish and German. For each language, participants were asked to indicate whether it was their first, second, third or fourth, respectively. Even though participants were expected to have grown up with French, exclusively, some did come from bilingual backgrounds, which were either French-English, French-other, English-other, or other-other. Except for the last option, which appeared rather improbable for the targeted population, the other possibilities were covered by alternative options in answering Q1 and Q2. That is, instead of assigning a number,

---

[32] The data set may thus be used for future investigations into Spanish-based transfer, when participants were presented with test items of Romance origin. In particular, three out of these six Latinate borrowings, precisely *Tasse*, *Möbel* and *Meer*, could equally be related to a Spanish counterpart: *taza*, *muebla* and *mar*, respectively.

participants could indicate which other language they had learned during childhood alongside French, or English, respectively. Questions 3 and 4 only required participants to indicate where Spanish and German were situated on this chronological acquisition scale. On the basis of the information thus acquired, we only included participants who had learned French during childhood, and English as their second language, as well as Spanish and German as their third and fourth languages, respectively[33]. In case the language indicated as the first in terms of order of acquisition was not the dominant language in the immediate family, the fifth question of this section was included to further specify the language(s) spoken at home with different family members, including the mother, the father, the grandparents and siblings.

As stated above, information about the participants' respective learning experiences of English, Spanish and German were gathered in the following sections, each of which consisted of the same five questions numbered continuously. Thus questions 6 to 10 focused on English, questions 11 to 15 on Spanish, and questions 16 to 20 on German. The first question of each section (Q6, Q11, Q16) was concerned with the age at which participants had started to learn the language. In the case of English, they also had to specify whether they had been attending a regular or an intensive English course at secondary level, given that this particularity may provide further insights as to the learners' individual capacity to make use of English vocabulary to understand new words in German.

The following question of each section (Q7, Q12, Q17) referred to the number of completed years, or – if less than a year – the approximate number of hours of formal instruction the participant had received in the given language. This question was included

---

[33] Note that this was the only place in our protocol where the numbering of languages referred to a purely chronological order of acquisition, i.e., where each language was assigned a number according to the point in time when a participant had first encountered and started to learn the given language. This was only interesting at this particular moment, since we wanted to be sure that our participants had comparable language backgrounds. Whether they called Spanish or German their "third" language did not make a difference as to the label L3, which was used for German throughout our study, given that the latter merely referred to the target language in our study, which participants were currently learning after having previously learned at least one other L2. Thus, in line with Hammarberg's (2010) typology, even subjects who had had experience with Spanish outside their classroom and/or learned Spanish before first encountering German, were included in the study and referred to as German L3 learners, just like those who had simultaneously started to learn Spanish and German at *cégep*.

in order to assure a more exact description of our participants, on the one hand, and to corroborate or to specify the previous answer indicating the age when they started to learn the given language, namely because language acquisition can take place in a natural or in a formal setting. Thus, it could have been the case that a participant had not received any formal instruction in German before entering the *cégep*, but that they considered they had started to learn that language at an earlier stage, or vice versa. In both cases, they were excluded from the study, given that they would probably have been familiar with the German test items.

The following question of each section (Q8, Q13, Q18) pertained to the amount of time that participants had previously spent in four specific linguistic environments, namely in a country, a family, work place, and/or in a school where the given language was spoken. This question was adapted slightly from Bayona (2009) insofar as we added the unit 'weeks' in addition to 'months' and 'years' and spent in the given language environment. This seemed important given that shorter trips might also influence the degree to which the words of a given language become more easily accessible and thus available as a source of transfer. Finally, the total amount of time spent in English-speaking environments was used as a secondary measure of Exposure to English, one of the potential predictors of positive transfer.

The last two questions of each of the three language-specific sections (Q9/Q10, Q14/Q15, Q19/Q20) pertained to the degree of interest in the given target language and culture, on a scale of five qualitative descriptors ("très faible", "faible", "moyen", "fort" and "très fort"). This affective dimension was included since learners with a higher interest in a given language and/or its culture may be more inclined to refer to that language when searching their lexicon for cues. Likewise, a higher degree of interest in German may be shown to relate to higher transfer rates, since these participants are presumably more inclined to want to understand the L3 on the basis of any sources available to them.

As for the measure of current exposure to English, the present protocol deviated consistently from the LEAP-Q. In line with Ranta and Meckelborg's (2013) reflections on

the concept, we adopted a broader view of L2 (or Lx) exposure in the sense that it included input, output as well as interaction in the given language. In fact, different types of contact (oral/written, virtual/real, receptive/productive) with a language or speakers of that language in different kinds of situations of everyday life can be expected to influence language activation. Marian et al. (2007) suggested six dimensions of exposure, namely (a) interacting with friends, (b) interacting with family, (c) language lab/self-instruction, (d) listening to music/the radio, (e) watching TV, and (f) reading, and asked participants to indicate their respective exposure levels on a scale from 1 to 10. However, a number of other activities such as playing video games or chatting on the Internet, in which the use of a non-native language (mostly English) seems considerably more probable for young adults, were not stated in their protocol. In our questionnaire, the last question (Q21) addressed language exposure in a table including ten dimensions of language use, for which participants were asked to indicate frequencies ranging from "daily" (Q = quotidiennement) to "never" (J = jamais) in each of their languages, including their native language French. This conception of bi- or multilingual usage was adopted from Grosjean (2014), who states that "bilinguals usually acquire and use their languages for different purposes, in different domains of life, with different people" (p. 3). Even if our participants did not grow up with two languages simultaneously, they were considered multilingual users, in the sense that their non-native languages, for the time being mostly English, would serve different purposes, in different domains of life, and with different people. Hence, it was judged appropriate to include all four languages in the grid and furthermore give participants the chance to add activities that might not be covered in the list, for which four additional spaces were left in the table. As for the measurement of the participants' relative exposure levels, the respective frequencies for each activity in the given language were quantified from 0 ('never') to 4 ('daily') and added up to a total score.

For the purpose of the present study, the data extracted from the activity grid described above (Q21) were subsumed under the variable named 'L2 exposure'. However, when the results were finally entered in a spreadsheet destined for statistical analyses, this primary exposure variable was labeled 'frequency of use' in order to be easily distinguished from a secondary type of exposure data, namely the questions pertaining to the amount of time

spent in different linguistic environments (Q8, Q13, Q18). As discussed along with the presentation of the descriptive results for each of the targeted variables (Chapter 5), the total number of weeks spent in different English-speaking environments was finally grouped into smaller time units (see section 5.3.2.2) to analyze this secondary exposure variable as a means of data triangulation (see also the *Coding procedures* represented in Table 4, section 4.5.1).

4.3.1.3 Validity and reliability of the background questionnaire

As mentioned above, the original version of the LEAP-Q (Marian et al., 2007) has been validated with respect to its underlying constructs such as language proficiency, language maintenance, or late language learning. In addition, independent proficiency measures have been used to validate the self-reported ratings, which proved to be consistent. On this basis, Bayona (2009) was able to justify the use of a modified version of the LEAP-Q to collect data on the proficiency levels in the different languages of her participants and use this measure as a predictor for transfer rates and types. In our study, however, the nature of the main task by which we were able to observe instances of (positive) transfer differed in a substantial way from Bayona's. While her participants were producing compositions in their L3 (Spanish) whose anomalous structures (classified into lexical, morphological and structural errors) were analyzed for crosslinguistic influences from different background languages, the participants in our study were faced with unknown words in the target language, exclusively. No underlying grammatical structures were possibly transferable from other languages in order to make sense of those items. All the items in our translation task were nouns in the singular, i.e., single-morpheme units, many of which had orthographic and/or phonological similarities with lexical material in different source languages. Therefore, the kind of source language proficiency that may be taken as an indicator for increased transfer rates was, above all, vocabulary knowledge in English. Hence, instead of including self-rated proficiency levels in the questionnaire, we chose an independent measure of English language proficiency, which focused on three receptive skills, precisely grammar, vocabulary and reading comprehension (see section 4.3.5).

4.3.2   Word translation task

4.3.2.1 Rationale for the word translation task

In order to collect data on positive transfer from English (L2) to German (L3), the main study consisted of an elicited task type, namely a *word translation task* from German into French, the participants' first language (L1). Test items included a number of unknown words in the target language (TL) German, many of which bore formal and semantic similarities[34] with English. As indicated in section 4.2, participants who reported prior knowledge of German through travel, German-speaking friends or multilingual households in the background questionnaire, were excluded from the study, in order to control for the level of German vocabulary knowledge at the moment of data collection. At early stages of foreign language learning in a formal setting, it was fairly predictable which vocabulary items had already been acquired. Since it was expected that the majority of the selected test items were unknown to the participants, potential lexical cues would have to be drawn from a source other than the TL itself.

Other researchers have designed translation tasks to elicit data on positive lexical transfer. Gibson and Hufeisen (2003) as well as Rast (2010), have used text translation tasks from an unknown language into the participants' mother tongue. Both studies revealed lexical transfer from other languages and mostly from those, which were typologically related to the given target. Even though our participants had already started to learn German in a formal setting and therefore had a minimal knowledge of the language at the time of data collection, they were expected to make use of their prior knowledge of English and/or other languages when confronted with new vocabulary in their L3. However, the reason why we chose a word translation task instead of a text is that the latter would in itself have to be considered a metalinguistic task, since reading comprehension clearly "requires the individual to think about the linguistic nature of the message: to attend to and reflect on the structural features of language" (Malakoff, 1992). In other words, when translating a text, learners do not solely have to draw on lexis but on grammar, syntax, logic, etc. In a more recent study on the advantage of bilinguals over monolinguals in learning subsequent languages, it was in fact a translation task using text from an unknown language that was

---

[34] The particular similarity relations, on which the selection of test items was based will be discussed below.

used to indicate levels of metalinguistic awareness in the different groups of participants (Kuile et al., 2011).

Therefore, it was expected that a word translation task was less likely to interfere with the measure of MLA itself, i.e., the THAM-3 (see above), since it was mostly the activation of what Ringbom called "potential vocabulary" (Ringbom, 1992, p. 88) that could be thought to help learners to make lexical inferences. This is reflected in Banta's (1981) statement on the common sources of German and English vocabulary:

> Fortunately, learning German vocabulary can be facilitated by using the English that we already know. A number of words will be immediately familiar: *Hunger, Mann, Butter, lernen, in, hier, oft, intelligent* will probably be understood at first sight or sound, although one is sometimes astonished at the inability of some students to recognize even such obvious cognates and common borrowings. Or is it inability? Is it perhaps mistrust? Do we make it sufficiently clear to our students that German and English are close relatives? (p. 129)

Within the framework of the present study, the goal was not to make it clear to the participants where to look for similarities but to have them find the cues by themselves in order to allow an identification of the factors that may or may not have led them to establish these links.

4.3.2.2 Description of the word translation task

The task took the form of a vocabulary test, namely a list of German nouns to be translated into French. The selected items (see Appendix C) were first presented in isolation, followed by a German sentence in which the word appeared. Even though this part may seem to contradict choice of the task (i.e., not to use a text but individual lexical items to be translated), we opted for a contextualized example accompanying each of the isolated items in order for the participants to consider them as a useful part of the target language. The reason for this was related to an anticipated reaction of learners faced with a task that required them to guess the meaning of words they do not know, which could have led to frustration (Friel & Kennison, 2001, p. 258). Furthermore, in most learning situations, be it in a natural or in a formal setting, vocabulary is not encountered out of context. Traditional

approaches to vocabulary learning and teaching, in which decontextualized vocabulary lists were commonly used, have gradually given way to communicative and task-based language programs (Read, 2004; Zimmerman, 1997). Thus, even when learners are confronted with unknown words in a classroom, these are most frequently related to some kind of communicative situation. Presenting the words within a sentence was intended to invoke the practical goal of vocabulary acquisition that is to enable learners to communicate. Naturally, participants were expected to try to make sense of the unknown word in looking for cues in the sentence. Even if the participants were provided sentence contexts for each test item, the sentences were constructed so as not to allow them to directly infer the words' meanings, given their restricted level of German proficiency. For example, when reading "Ein *Schiff* fährt über den Fluss", the learner would have to know the word *Fluss* (English: *river*) and be familiar with the verbal construction *fahren über* (English: *to cross*) to infer the meaning of *ship* based on the context, exclusively. In a much simpler example, when faced with the sentence "Wir haben einen *Sohn*", learners might correctly understand the context as "we have a [something]" but would still need to establish a link to the English cognate *son* in order to translate item correctly. Essentially, our goal was not to include semantic cues such as *Tochter* (English: *daughter*) in combination with *Sohn*, for example. At the same time, the sentences could be referred to as a means of justification. More precisely, even if this was not necessary to transfer individual items based on their English cognate, a cross-check with the sentence could help participants make the final decision. This can be illustrated with the example stated above: Whereas "we have a *son*" is a sensible thing to say, "we have a *sun*" or "we have a zone" might raise conceptual questions, even if there are formal similarities to be established between *Sohn* [zo:n] and *sun* or *zone*, respectively.[35] In any case, similarity relations to words from other languages have to be established in order to provide the correct answer. In other words, we can assume that correct translations are due to making the link between languages, especially English and German, unless the word was known before.

---

[35] Concrete examples of different kinds of sentence processing will be given in Chapter 5 where the results are presented and discussed.

Each item was accompanied by an auditory sample of its pronunciation, which was launched automatically as the item appeared on the screen. This was followed shortly thereafter by the appearance of the sentence including the item in question. Both the isolated item and the sentence remained on the screen for 30 seconds, during which time participants were asked to record their answers, namely the translation equivalent in French, as well as their verbalized thoughts on the translation process (see section 4.3.3.2 on TAPs). As research in foreign language vocabulary acquisition has shown, items from different languages that are written and/or pronounced alike, tend to be co-activated more easily if accessed phonetically (de Bot, 2004; Fitzpatrick, 2006; Meara, 2009), which explains the importance of providing learners with both orthographic and auditory samples of the target material, if crosslinguistic relations are to be established. However, to invoke the notion of *lexical access* to justify the use of a phonetic support seems rather unsuitable in this particular context, since the participants had not yet learned the majority of the TL items that they were meant to translate. Thus, when first encountered, these words would not have any semantic and conceptual representation attached to them. Therefore, in order to translate these words the process the participants had to engage in was one of cognate[36] recognition, rather than of lexical retrieval. Only when a similarity relation to an available item from another language could be established, was there an attribution of meaning. The additional auditory information was meant to facilitate those meaning-making strategies. Specifically, initial stage learners of German were not expected to have a representation of the appropriate pronunciations for the unknown words, even though they may have been taught, for instance, that the diphthong [au] is usually represented graphically as "au", in contrast to English "ou" or French "aou". Similarity relations between related languages such as English and German can often be established on both orthographic and phonological levels. This is the case of homophones such as *Haus* and *house*, both pronounced [haus], or homographs such as *Hand* [hant] and *hand* [hænd] in German and English, respectively. Hence, the identification of sound-based correspondences was useful when the orthography of the TL item was not transparent enough for the learner to associate it to a cognate from English. Researchers have confirmed this assumption in a similarity-rating task of German-English translation pairs by monolingual speakers of

---

[36] The notion of cognate will be discussed in more detail below.

English, some of whom were provided with a pronunciation guide while others only had access to written samples. Significant differences between the two conditions could be identified for German items that included German-specific characters such as *ß* or Umlaut-vowels such as *ä*, *ö* or *ü*. The authors explained inferred that the participants, who did not know any German, could not possibly infer the appropriate phonetic information (Friel & Kennison, 2001, p. 258). Even though the participants in the present study already had a small basis in German, the auditory samples were expected to facilitate the establishment of similarity relations to translation equivalents from other languages, in this case mostly from English. In turn, this was supposed to further prevent the relative frustration related to a task that exceeded the demands of a regular vocabulary test because the items had not been studied before.

4.3.2.3 Choice of test items for the word translation task

Given that the task required participants to infer meaning based formal correspondences between words from different languages it involved the recognition of *cognates*. Since this term has often been used to denote a variety of different similarity relations, we will briefly summarize those definitions before stating the one adopted for the present study. Whereas cognates are commonly understood as words from different languages "which share both semantic and phonological information" (Burton, 2013, p. 50), some researchers also include the orthographic level, thus leading to definitions such as "translation equivalents […] similar in appearance and sound" (Friel & Kennison, 2001, p. 249) or slightly more flexible ones such as "cognate words share (parts of) their orthographic and/or phonological form with their translations […]" (de Groot & Keijzer, 2000, p. 3). The particularity of these definitions is that they do not exclude borrowings, where such formal correspondences between translation equivalents are often most striking yet not based on typological proximity. Such would be the case for English-German word pairs such as *revolution-Revolution*, *reality-Realität* or *president-Präsident*; these translation equivalents are nearly identical in shape and sound, yet this is due to the fact that they were borrowed from French or Latin, and not to the fact that English and German are genetically related within West Germanic branch. To consider such Latin-based borrowings as English-German cognates would imply the same relation to the French (or Spanish, or Italian, or

even Polish, etc.) counterpart, in which case the typological, or rather the genetic, relationship between English and German would not play any considerable role for the acquisition of the given item in the TL. Hence, the kind of cognates of interest for our study are the ones that have "common etymological roots *and* similar forms and meanings" (Dressler et al., 2011, p. 243, italics ours). More precisely, the English-German translation equivalents from which positive transfer was expected were "pairs of words that show sound-meaning correspondences indicating their historical relationship" (Banta, 1981, p. 129). Finally, this relationship must be anchored in the Germanic-based etymological roots of the given word pairs, such that the same similarity relations to the French or Spanish translation equivalents could not be established simultaneously.

The test items were selected from a large-scale study (Friel & Kennison, 2001), in which monolingual speakers of English were to identify the respective *cognate*, *false-cognate* or *non-cognate* status of 569 German nouns, selected from German education textbooks and from various issues of the German news-magazine *Der Spiegel*. Two hundred and fifty undergraduate students from an American university participated in the study. The tasks that served as a basis for this identification were a similarity-rating task of German-English translation pairs, on the one hand, and a translation elicitation-task from German to English, on the other. Each was administered to a distinct group of one hundred participants. Task results showed a highly significant positive relationship between similarity ratings and translation accuracy rates. With reference to Kroll and Stewart (1994), cognate status was established for word pairs that were translated correctly more than 50% of the time. Moreover, apart from the degree of similarity of the word pairs, which was rated on a 7-point scale, and their "cognate" status, two other important word characteristics were evaluated. Namely, the remaining fifty participants were engaged in an imageability rating-task for the English translation equivalents, a measure which seems to be largely consistent with ratings of word *concreteness* (e.g., de Groot & Keijzer, 2000; Qasem & Foote, 2010). And finally, by reference to selected databases for each language, printed word frequency was identified for each item. Thus, using items from Friel and Kennison's lists (2001, pp. 262–274) for the present study made it possible to examine tendencies with respect to these other word characteristics (see section 5.1.3 in Chapter 5).

The word pairs classified as "cognates" by Friel and Kennsion (2001) included Latin, Greek and French-based borrowings, such as, for instance, *Aktion-action*, *Energie-energy* and *Thema-theme*, which are also nearly identical in French. Out of 112 "cognates" thus identified, we first extracted all those German items that were etymologically related to English, but not borrowed from Latin, Greek or French and thus do not have identical translation equivalents in French (see Table 2 below). From the remaining items, we further discarded all those with cognates in French, since it would otherwise have been possible to guess the correct meaning by reference to French, and not to English. Furthermore, German-English homographs such *Arm-arm*, *Fall-fall* or *Winter-winter* discarded since participants could have interpreted them as English words in the first place, and would possibly not have engaged in any comprehension strategy whatsoever.

The second step involved the analysis of Friel and Kennison's (2001) list of word pairs identified as "non-cognates", due to the fact that their participants had not reached the 50 % line of correct guesses in the translation task. However, the differences between their participants and those who took part in our study must be taken into consideration. Namely, the fact that the majority of the monolingual speakers of English were able to translate a number of words from an unknown language correctly suggests that speakers of two or more languages, including English, should be able to obtain similar if not superior results with respect to translation accuracy based on similarity relations. This assumption was based on findings in SLA and TLA research, indicating that foreign language learning experience provides bilingual and multilingual learners with the cognitive flexibility that is needed to perform a perpetual search across their languages in order to meet the demands of the given linguistic environment (Cook, 2002; Cummins, 1991; Jessner, 2006). We extracted a number of word pairs displaying roughly the same characteristics as those extracted from the cognate list in as far as they were closely related and thus resembled their English counterparts both orthographically and phonologically. Finally, we selected all those Germanic-based word pairs that were classified by at least 4 out of 7 points for

their degree of similarity in at least one of the two conditions.[37] At the end, the list included ten (10) items from the cognate list and eighteen (18) that displayed a high degree of similarity, despite their identification as "non-cognates". The participants of our study who were able to establish the link between the unknown target item and its English translation equivalent were therefore expected to be able to translate a total of 28 test items into French.

Finally, in order to supplement our translation task with items that were not similar to English, fourteen (14) more German words were added. Six (6) of these were Latin-based borrowings from French and thus resembled their French counterparts, while the corresponding term in English did not show any similarity. Given that we could only identify three French borrowings from Friel and Kennisons's study that did not share any similarities with English, namely *Tasse*, *Krawatte*, and *Fenster*, in contrast to, for example, *Tante* (English: *aunt*) or *Regel* (English: *rule*), we added three more comparable items, in order to attain a higher ratio with respect to test items where the source language to provide a lexical cue was French, and not English. Otherwise, participants could have taken it for granted that they were supposed to draw on their English vocabulary in order to make sense of the unknown words in German. Finally, eight (8) more items were chosen from the "non-cognate" list that displayed neither any similarity to English nor to French, so that one third (14/42) of the test item could not be guessed on the basis of English translation equivalents.

---

[37] Two exceptions were made in the case of *Regen* (*rain*) and *Brust* (*breast* or *chest*). The former was included despite a similarity rating of only 3.04 in the pronunciation guide condition. For the latter, the translation equivalent provided to Friel and Kennison's participants was *chest*, instead of *breast*. Hence, the similarity rating was very low (2.12). However, both translations are acceptable and it seemed plausible to recognize the similarity relation between *Brust* and *breast*.

Table 2. *Translation Task – Test items*

Similarity to English		Similarity to French/Spanish		No similarity
German	English translation	German	French/Spanish translation	German
Apfel [ˈapfl̩]	apple	Tasse [ˈtasə]	tasse/taza	Messer [ˈmɛsɐ]
Feld [felt]	field	Krawatte [kraˈvatə]	cravate	Spiegel [ˈʃpiːgl̩]
Fisch [fɪʃ]	fish	Fenster [ˈfɛnstɐ]	fenêtre	Zeitung [ˈtsaitʊŋ]
Freund [ˈfrɔynt]	friend	Ferien [ˈfeːrɪ̯ən]	(jour) férié/fería	Berg [bɛrk]
Grund [ˈgrʊnt]	ground	Möbel [ˈmøːbl̩]	meuble/muebla	Lied [liːt]
Haar [haːɐ̯]	hair	Meer [ˈmeːɐ̯]	mer/mar	Pferd [pfeːɐ̯t]
Knie [kniː]	knee			Kleid [klait]
Socke [ˈzɔkə]	sock			Luft [lʊft]
Sohn [zoːn]	son			
Sturm [ʃtʊrm]	storm			
Stein [ʃtain]	stone			
Schiff [ʃɪf]	ship			
Regen [ˈreːgn̩]	rain			
Maus [ˈmaus]	mouse			
Milch [ˈmɪlç]	milk			
Herz [hɛrts]	heart			
Brust [brʊst]	breast			
Blut [bluːt]	blood			
Bäckerei [bɛkəˈrai]	bakery			
Gast [ˈgast]	guest			
Kuh [kuː]	cow			
Kirche [ˈkɪrçə]	church			
Wasser [ˈvasɐ]	water			
Wetter [ˈvɛtɐ]	weather			
Sonne [ˈzɔnə]	sun			
Buch [buːx]	book			
Nachbar [ˈnaxbaːɐ̯]	neighbor			
Licht [lɪçt]	light			

4.3.3   Think-aloud protocols

4.3.3.1 Rationale for the think-aloud protocols

In order to highlight the cognitive process that the participants engaged in, they were asked to produce *think-aloud protocols* while completing the translation task. More specifically,

oral recordings, in which their individual meaning-making strategies were verbalized, allowed a verification of whether or not positive transfer from English was reported as the reason for correct translations from German into French. As mentioned in the previous chapter, the question of how to clearly identify transfer phenomena remains problematic, even if researchers have access to production data of deviant target TL use. In the case of positive transfer, where L2 or L3 learners use their TL appropriately, this identification problem applies even more. In the present study, it was the recognition of formal similarities between words from different languages that would lead participants to produce correct translations. However, instead of establishing such connections to other languages, they could also have made a lucky guess and give the same correct answer. Hence, the necessity of collecting introspective data arose from a more general observation, namely that "understanding the source of second language production is problematic because often there are multiple explanations for production phenomena that can only be assessed by exploring the process phenomena'' (Gass & Mackey, 2000, p. 19).

Think-aloud protocols (TAPs) have mainly been used in cognitive psychology to collect introspective data during or shortly after a given problem-solving task. The first major contribution to a research methodology using TAPs goes back to the 1980s (Ericsson & Simon, 1993 [1984]). Since then, they have been used to investigate cognitive processes in other fields such as clinical health studies (e.g., Fonteyn, Kuipers, & Grobe, 1993), translation research (e.g., Bernardini, 2001), as well as second language acquisition (SLA) (e.g., Faerch & Kasper, 1987). As for the theoretical baseline, Ericsson and Simon's (1993 [1984]) framework was based on a model of information processing which places short-term memory in the center of so-called "concurrent probing" techniques, that is the verbalization of thoughts, while participants are working on the given task, as opposed to "retrospective probing" which takes place after task completion (pp. 11-12). The particularity of short-term memory is that the information one has only just attended to can be accessed directly and thus easily reported in verbal form. Even though both concurrent and retrospective protocols are subsumed under the larger category of 'verbal reports', the term "think-aloud protocol" most commonly refers to introspective data that reflect on-line processing. Depending on the research orientation, TAPs may or may not involve what

Bowles (2010, p. 14) refers to as "metacognitive" verbalizations, i.e., where participants are asked to give explanations and justifications for their reasoning. Generally speaking, Ericsson and Simon's (1993 [1984]) claim is that TAPs do not influence the task results as long as the verbalization, metacognitive or not, represents the cognitive processes that participants would normally have engaged in when solving the given problem.

However, the main challenges to the use of TAPs have been posed with respect to its validity regarding the nature of the cognitive processes they are thought to mirror. Specifically, the verbalization is an additional task that in itself may alter these processes. Observations of this kind have been made with respect to problem-solving tasks of non-verbal nature, as conducted in cognitive psychology, and, more recently, in verbal tasks in the field of SLA (as discussed in Bowles, 2010). However, while TAPs tend to slow down the problem solving processes the participants are currently engaged in, the degree of reactivity that such on-line verbalizations may cause for accuracy seems to be strongly task-related. As Bowles (2010) points out in her literature review of non-SLA studies where the effects of verbalization on the outcome of different types of problem-solving tasks were tested, a general tendency can be observed: First of all, non-metacognitive verbalizations do not seem to influence task results whereas metacognitive ones sometimes do, either hindering or improving task accuracy. However, in most of these cases, it could also have been the complexity of the task itself that may have caused reactivity. More specifically, only if a task was cognitively highly demanding for the participants, would verbalization tend to interfere with the task results (Bowles, 2010, p. 64). This indicates that verbal protocols, especially those that are metacognitive in nature, i.e., that demand explicit explanations for decision-making, are likely to hinder the performance in highly complex problem-solving tasks, namely because the cognitive load of these combined measures is too high.

However, as Bowles (2010) also points out, these findings cannot necessarily be taken as good indicators regarding the use of TAPs in SLA, given that only very few of the tasks used in cognitive psychology addressed verbal behavior. Hence, the researcher undertook an exploratory meta-analysis of nine SLA studies in which the use of TAPs was tested with

respect to reactivity on different comprehension and production measures in a second language. However, due to the variability in the respective research designs with respect to numerous aspects such as task type (e.g., writing vs. reading, item vs. system learning), learner level (beginner vs. advanced) or the type of verbalization (metacognitive vs. non-metacognitive), it was impossible to clearly indicate whether or not TAPs influence task performance in SLA, merely by looking at the results. In sum, Bowles' (2010) meta-analysis revealed small effect sizes, both positive and negative, for all measures of accuracy, ranging from text comprehension to receptive and productive form learning. As far as the distinction between metacognitive vs. non-metacognitive verbalizations was concerned, the latter was found to have a small facilitative effect on comprehension across the different studies, whereas no such effect could be determined for metacognitive reports, given that only one of the studies had investigated the use of that protocol type with a measure of comprehension. However, it should be noted that the findings of that particular study (Bowles & Leow, 2005) did reveal a significant (but negative) effect of metacognitive TAPs on text comprehension. However, there was no evidence for this effect in the meta-analysis, because "there were insufficient data to calculate an average effect size for metacognitive verbalization on this measure" (Bowles, 2010, p. 93). In comparison, small to medium sized detrimental effects were registered for both types of TAPs on productive and receptive form learning, respectively (p. 106). In other words, it seems to depend on the type of L2 performance targeted by the given task whether a specific kind of verbal report influences that performance. Bowles' observations further suggest that other factors such as the language to be used (L1 or L2) in the verbal report and in the task, respectively, as well as the level of L2 proficiency may also influence the way in which TAPs interfere with task performance. Notwithstanding, the overall effect sizes for each of the observed factors that could, at least partly, be held responsible for the reactivity of TAPs in SLA studies, were very small. Furthermore, in her conclusive remarks, Bowles (2010) stresses the fact that the 95 % confidence intervals overlapped the zero value in nearly all of the effect size calculations (p. 138), thus indicating that the use of TAPs, metacognitive or not, does not seem to compromise the internal validity of the performance measures used in these SLA studies. Hence, she concludes that "verbal reports can reliably

be used as a data collection tool" (p. 138), thus corroborating Ericsson and Simon's (1993 [1984]) claims.

It should be noted, however that the reactivity of TAPs has only been investigated with a small variety of L2 performance measures, none of which actually resembles the word translation task that our participants were asked to engage in. We know of at least one study in the field of TLA in which TAPs were used to collect data on the positive influence from supporter languages, namely Jessner (1999, as discussed above). The results revealed how experienced learners made use of all their languages in order to produce L3 texts. It was shown that they were consciously looking for similarities between Italian, German and English when in doubt about specific features of the target language. As stated by the author, these data "provide evidence of the metalinguistic thinking involving the usage of all three, typologically closely related, languages in the process of academic writing" (Jessner, 1999, p. 205). If the author stressed the metalinguistic dimension of the task, this was not due to the type of verbal report that was targeted, since participants were merely asked to think out loud whatever went through their heads during the writing process (p. 204). As far as reactivity is concerned, no observations could be made given that there was no silent control group. At the same time, reactivity was not an issue in Jessner's protocol, given that no post-test measure was carried out to investigate the effects of a certain kind of treatment. It may be assumed, however, that verbalization could have influenced the degree to which the participants actively used their prior language knowledge to find appropriate words for their writings in the L3, namely by making their cognitive processes explicit, even if possible effects of verbalization on academic writing tasks have, to our knowledge, not been addressed in the literature. Based on the assumption that verbalizations do not appear to influence task performance if they reflect the kind of reasoning that participants would naturally engage in while solving the given problem (see above), Jessner's (1999) findings may be taken to mirror fairly accurately the ways in which advanced L3 learners resort to their other languages to fill the gaps in TL knowledge when writing a text in their L3 without a dictionary. It must be acknowledged, however, that the mere fact of verbalizing may facilitate cognitive processing. The section on the limitations of the present study with regard to the think-aloud procedure (6.3.2) includes a discussion on the extent to

which the interdependence between speech and thought makes it difficult to consider verbalization as a neutral methodology (e.g., Swain, 2006b).

4.3.3.2 Description of the think-aloud protocols

In our word translation task, the participants were presented with a series of items whose orthographic and phonological forms may or may not give rise to associations with words from other languages, which, in turn, may or may not lead to correct translations. Far from the complexity of language production tasks, where learners have to make flexible use of various facets of the linguistic (including lexical) knowledge they possess, our task can hardly even be considered as a measure of "comprehension" of new vocabulary, if the latter term is meant to encompass all aspects of word knowledge. Rather, our translation task aimed at measuring a kind of lexical inferencing based on formal similarities, i.e., the recognition of cognates.

The participants used the computers of each school's language laboratory to complete the task individually, using the program DmDX, which provided an interface that recorded their answers given orally in a time frame of 30 seconds, following the cross-modal stimulus, namely the written word, both isolated and in sentence context, accompanied by a sound sample (of the word, not the sentence). Apart from the respective translation equivalents that seemed appropriate for each item, the participants were asked to verbalize the mental process by which they arrived at a result for each answer. However, in order not to orient their reasoning or to put an additional cognitive load on them, they were not explicitly asked to justify their answers, but simply to verbalize everything that came to mind while trying to translate the test items. The only way to guess an appropriate translation was by association, unless, of course, a given word had already been encountered at some point. The verbal protocols revealed both of these possibilities. Namely, if a German word was recognized because it had already been learned – in class or somewhere else – this could become apparent in the protocol. If, however, the word was unknown, associations were triggered by sounds and/or letter strings that evoked referents in the different languages known to the participants. The word "Luft" ('air'), for instance, which has no translation equivalent in English, French or Spanish that could serve as an

orthographic or a phonological cue, was sometimes associated with "lift" or "loft", which are also used in Québec French. On the other hand, correct translations could be provided based on lexical cues from English such as, for instance, "souris" for "Maus" based on English "mouse", and thus reveal positive transfer from English, whereas the cue could also have been provided by the sentence displayed with the item. In the case of "Maus", the words "klein" ('small') and "Tier" ('animal') may have served as basis for interpretation, in which case the translation strategy was not transfer from English, but lexico-syntactic inferencing from German. However, given the participants' level of German and the simplicity of the task, the expected type of inferencing was on the basis of formal similarities with words from other languages, namely where such orthographic and/or phonological correspondences could be established. However, an analysis of basic grammatical aspects was also expected to help participants deduce the right answer. More precisely, the syntactic information in the support sentences could remind them that they were looking for nouns, namely because they were preceded by determiners or adjectives. Thus, even though the lexical information provided in the sentences was not of much help, participants who validated their answers with the structure of the respective support sentences clearly had better chances of finding (or confirming) the right answer (see Chapter 5).

In order to validate the accuracy of our instructions, as well as the length of exposure to each item, we conducted a pilot study with five students from *Cégep Beauce-Appalaches* who were at the end of their first year of learning German in a formal setting. The collected material of this test run (21 minutes of oral recording for each participant) also provided some crucial insights into the kinds of reasoning we could expect to find in the main study. Initially, we could confirm that the instructions that were read out loud by the researcher while the students could read them on their screen were clear. They understood that they had to translate the words orally and that they were to think aloud while searching for the right translation. This was made obvious in the verbal protocols. Even though they had not been advised to do so, they naturally tried to find formal correspondences to words from other languages to infer the meaning of unknown vocabulary items. However, given the

time of year (March, 2014), a number of words were already known to the students who participated in the test run.

As an overall observation regarding these preliminary results, a high degree of variability was expected for this test when administered at a larger scale, given the notable qualitative differences singled out in the test run, which was carried out with a very small number of students at similar proficiency levels in English and in German. Moreover, in a short conversation with the researcher following the test run, all four participants averred their willingness to learn new words in German, and felt this exercise helped them understand unknown vocabulary or reaffirm what they had already been introduced to in class.[38] Such a positive attitude towards the language and the test itself can certainly be taken as a favorable condition for participants to search thoroughly for possible cues in order to make the most of this experience. Hence, if these four students were also comparable in terms of attitudinal and motivational factors, the qualitative differences in depth of analysis are probably due to other individual factors, most likely including metalinguistic abilities. What was further revealed by the students' comments was that the time frame of 30 seconds per test item was considered appropriate. Even though two students asserted they would have liked to have more time for some of the unknown items, there were only two cases out of 168 (4 x 42 test items) for which no answer was recorded. Instead, they said that they did not know the answer. From this, however, we could not deduce that it was actually due to the limited time frame that they did not find an answer, given that their verbalizations were not cut after 30 seconds. Only for those items that the participants had no doubt they knew already, they perceived the time as too long.

We thus concluded that the 30 seconds allocated for each item in our DmDX script would be retained for the main phase of data collection. Last but not least, we noticed that the physical proximity between the students in the language laboratory, which made it possible

---

[38] These comments that followed the administration of the pilot test indicate that it may have been the act of speaking during the translation task that helped these students to make sense of the German vocabulary. As will be discussed with respect to the limitations of the present study, future investigations in which TAPs are used to collect data on meaning-making strategies would benefit from follow-up interviews or other means of data triangulation to examine the effect of verbalization on the cognitive processes that learners engage in during the task.

for them to overhear each other's answers to the same test items, influenced their task performance in a fundamental way. When they heard someone else's answer, they would either use that as a hint for their own reasoning or get confused. In any case, we realized that this immediate influence could distort our results significantly. Two conclusions were drawn from this observation. First, we planned to leave as much space as possible between our participants of the large-scale study. Second, instead of fixing the sequence for each individual test interface to be identical, we adapted the script to have the order of test items randomized. Consequently, participants were not working on the same items at the same time and were thus less likely to influence one another, even if their respective comments could still be overheard.

4.3.3.3 Validity and reliability of the think-aloud protocols

The controversy about the validity of research designs including verbal reports in general, and think-aloud protocols in particular, stems from a body of research displaying a high variability in the possible factors that might be responsible for altering the outcome of a given task. A closer look at a number of these studies in the fields of cognitive psychology and SLA has revealed that the use of TAPs during a given verbal or non-verbal problem-solving task should only minimally affect the task results, as long as the verbalizations that subjects are asked to produce reflect the line of thoughts that they would normally have engaged in during the given task (see above).

According to Bernardini (2001), there are some major implications of Ericsson and Simon's (1993 [1984]) information processing model for the use of TAPs in translation research[39]. On the one hand, she underlined the importance of using concurrent instead of retrospective probing if the goal of the given study is to "exhaustively reflect mental states" (for tasks with a duration of more than 10 seconds), given that the information to be retrieved from long-term memory after the completion of such a task is likely to be incomplete and/or subject to interpretation (Bernardini, 2001, p. 243). On the other hand,

---

[39] Even if the present study is not situated in the field of translation research, Bernardini's (2001) considerations refer to the use of verbal reports as a means of highlighting "translation strategies" in a larger sense. Namely, she addressed a body of research in SLA (pp. 244-246), before concentrating on the more specific processes involved in text-translation.

she pointed to the environmental validity of TAPs which can only be assured if the participant is not restricted by any of the social norms that may influence the information a person is attending to when interacting with others (p. 243). In other words, TAPs should be used with tasks that are solved individually. A further consideration concerns automaticity in the completion of a task (or task type) that participants already have experience with, given that the latter require less conscious control of the mental processes involved and will thus be less easily available for verbalization. For the context of our study, these observations may serve as a partial justification of using TAPs to collect data on positive transfer. Namely, the word translation task complemented by the introspective data took the participants 21 minutes to solve. Retrospective data might not have accounted for the on-line processes carried out during the task. It was also an individual task, and it was focused on spontaneous associations with word forms, i.e., information in short-term memory that had not necessarily been treated and stored. Given that the German words to be translated were mostly unknown to the participants, information to be retrieved from long-term memory was only related to lexical material from previously acquired languages. Yet only the associations that came to the participants' minds at that very moment could be used to infer meaning. Moreover, the recognition of test items they might have come across prior to taking the test, was also expected to surface in short-term memory, and might lead to verbalization of those items. Even if such recognition did not lead to correct translations, verbal reports could attest that the association leading to the given response was not related to one of the learner's background languages. The preliminary verbal data collected in our pilot test confirmed the usefulness of this introspective method to indicate when participants made use of their background languages. A further example should illustrate this point: *Kirche* (English: *church*) was correctly translated by two participants of the pilot test, one of whom mentioned how she had learned the word in a conversation class with the language assistant outside class, while the other concentrated on the syntactic structure of the support sentence, underlining the fact that it had to be a place, which apparently made her think it could be a church. Even if she did not explicitly mention the English translation equivalent *church*, there is reason to assume that she intuitively opted for the correct answer *église* on the basis of an association with the English counterpart *church*. Of course, it is impossible to argue with certainty that the answer was not due to pure chance.

However, given that initial-stage learners were shown to rely heavily on formal aspects of words to infer meaning (e.g., Ringbom, 1992, p. 89), explicit references to syntactic and semantic aspects of the supporter sentence are expected to represent additional cues to the more basic formal associations, which in turn may be conscious or unconscious. Hence, correct translations that were either explicitly or implicitly related to an English cognate were taken to reflect positive transfer from English, unless participants reported that they had already learned the word. In other words, without access to the reasoning behind them, all correct translations would have to be rated as positive transfer. Moreover, the German teacher confirmed that the word *Kirche* had not yet been seen in class. In sum, when test items were in fact encountered for the first time, the most probable source of association was formal similarity to English.

In conclusion, it should be kept in mind that the motivation to use introspective data in this study was twofold. First, the analysis of the TAPs allowed us to exclude instances of prior knowledge and thus to confirm that correct translations were either explicitly or implicitly based on associations with English. In this sense, the verbal protocols were the chief indicator of positive transfer in the present research.[40] Second, and more importantly, the introspective data served "to describe how [learners] put information together when given a problem to solve" (Fonteyn et al., 1993, p. 431). Thus, the qualitative analysis of the TAPs allowed us to shed light on the cognitive processes that were triggered during the translation of mostly unknown words in German (L3). As discussed with respect to the coding procedures in the *Data analysis* section (4.5), verbalizations were coded for transfer, on the one hand, and for levels of metalinguistic analysis, on the other.

4.3.4   Metalinguistic ability test

4.3.4.1 Rationale for the metalinguistic ability test

As discussed in the conceptual framework, the notion of metalinguistic awareness (MLA) has been delineated differently according to the diverse research orientations in the

---

[40] It must be acknowledged that a questionnaire targeting prior knowledge of the target items could have been administered prior to the translation task. Any items of which prior knowledge was attested could thus have been excluded from the task altogether. This aspect will discussed in more detail with respect to the limitations of the study (see section 6.3)

respective fields, ranging from developmental psychology to educational psycholinguistics (see above). Even within the more confined circle of applied linguists, definitions of the basic concept remain largely inconsistent, which, in turn, is reflected by a variety of data collection methods. To resume briefly what has already been said earlier, the most common instruments that have been used to measure MLA are grammaticality judgments and error correction tasks (e.g., Abu-Rabia & Sanitsky, 2010; Lasagabaster, 2001; Serrano, 2011). Interestingly, only a few researchers have combined both of these task types to provide a more complete picture of what they have defined as MLA (e.g., Gibson & Hufeisen, 2006; Lefrançois, 2000). Naturally, it depends on the conceptual framework adopted in a particular study whether a given instrument is considered appropriate to operationalize the concept. However, even researchers who seem to agree on Bialystok's definition regarding the two basic components of MLA, precisely "analysis of representational structures and control of selective attention" (Bialystok, 2001b, p. 177), the respective methods of data collection may still differ in substantial ways. Thus, Lefrançois (2000, p. 71) varies the tasks in order to address both aspects, namely by using grammaticality judgments of different sentence types as well as error detection and correction tasks of different grammatical structures. With reference to Gombert (1990), she further underlines that knowledge of grammatical rules (as assessed by grammaticality judgments, for instance) is an insufficient indicator for metalinguistic abilities, but that they must also be reflected by an appropriate application and explicitation of these rules.

> Comme le définit Gombert (1990), il s'agit de mesurer la capacité de raisonner sur la syntaxe et de contrôler l'usage des règles de grammaire. Rappelons qu'il ne suffit pas ici d'être capable de formuler une règle de grammaire, mais bien de l'utiliser dans un contexte langagier ; dans les formes les plus développées des capacités métasyntaxiques, il ne faut pas seulement savoir pouvoir juger si une phrase est grammaticale ou non, mais aussi savoir expliquer pourquoi elle l'est. (Lefrançois, 2000, p. 144)

Interestingly however, few of the employed tasks in her protocol require subjects to provide a justification for their grammaticality judgments and error corrections. Likewise, Gibson and Hufeisen (2006) draw on Bialystok's dichotomy in using two types of grammaticality judgment tasks, one of which is focused on attention control, by assessing the capacity to

ignore semantic errors in an on-line listening task while detecting grammatical errors, whereas the other consists in detecting and correcting the same type of errors in a post-hoc reading task, thus requiring more structural analysis (Gibson & Hufeisen, 2006, p. 142). Despite overlapping theoretical orientations between Lefrançois (2000) and Gibson and Hufeisen (2006), the dimension of rule extrapolation as a partial measure of MLA is not taken into consideration in the latter study.

To our knowledge, the most comprehensive test battery to measure MLA empirically was developed by Pinto and her colleagues (1999), whose focus was precisely on the distinction between linguistic and metalinguistic levels of analysis, the latter being reflected by justification tasks. These researchers designed their "metalinguistic ability test" (MAT) for three different age groups, namely for children from 4 to 6 (MAT-1), for children from 9 to 13 (MAT-2) and for adolescents and adults from the age of 16 (MAT-3). According to research findings which indicate an increased development of abstract thinking in adolescents (Pinto et al., 1999, p. 109), the MAT-3 also measures the comprehension of syntactico-semantic relations and of figurative language, along with grammaticality judgments. Each section of the test has a linguistic (L) and a metalinguistic (ML) dimension, according to the respective levels of analysis, as reflected by two types of questions. The authors stress the fact that the metalinguistic answers clearly reflect *explicitness*, *analysis* and *intentionality*, thus "justify the use of the term 'metalinguistic' in its strongest sense" (Pinto et al., 1999, p. 168). According to repeated test runs by the researchers themselves, the MAT has proven to be a reliable instrument (Pinto et al., 1999, p. 169), which has been confirmed in subsequent studies using this battery (e.g., El Euch, 2010, p. 30).

4.3.4.2 Description of the metalinguistic ability test

In search of an appropriate method to assess our participants' levels of MLA, we analyzed the MAT-3 regarding its applicability to the context of our study. The original version of the test is divided into three sections: *(I) Comprehension*, *(II) Acceptability* and *(III) Figurative language*. The first section consists of 8 items each of which consists of two sentences, exposing qualitative, temporal, morphological or spatio-temporal relations. As a

first step, linguistic questions (LQ) require subjects to analyze each set of sentences with respect to their meaning, namely whether they represent the "same" or a "different" type of relation. Secondly, metalinguistic questions (MLQ) call for a justification of these answers in expressing the underlying structural features or rules that led to their conclusions. Since this process reflects the participants' explicit knowledge, it is the justification part that represents the metalinguistic dimension of the test. The same distinction between linguistic and metalinguistic levels of analysis is made in the second section (*Acceptability*) in which participants have to identify and correct 15 errors in a text. For each error, the identification and correction represents the linguistic dimension while the degree of analysis in the justification of each answer is quantified to represent the metalinguistic score of this section. The test items in the third section (*Figurative Language*) display metaphorical expressions, advertising slogans and excerpts from poetic texts. This part is used to measure the capacity of interpretation when facing figurative language. More specifically, participants have to indicate which meanings can be attributed to certain words in the given expressions by analyzing the metaphorical or humoristic levels underlying the overt linguistic structure.

Whereas the linguistic answers can easily be quantified as accurate (1) or inaccurate (0), the metalinguistic answers must be interpreted qualitatively according to their depth of analysis. For each test item, the researchers provide a number of possible answers reflecting three distinct levels of analysis, which largely correspond to (0) insufficient, (1) partly justified and (2) wholly justified answers. Thus, ML-scores can be added up according to a 3-point scale for each item. Even though, technically speaking, it is solely the ML dimension of the test that reflects levels of metalinguistic awareness, most test items require the participants to meet the linguistic demands of the test in the first place in order to access the metalinguistic sphere in which they exhibit their explicit knowledge of the given phenomenon by justifying their answers. As shown by different researchers having used this tool to measure MLA, it was of highest interest to include the L scores in the protocol for comparative purposes. The first empirical results of the MAT-3, for instance, showed how most of the 18 to 19 year-old Italian high school students obtained very high L scores on all three sections of the test, while their ML scores were much more variable (Pinto et

al., 1999, p. 166). In a more recent study where bi- and trilingual learners' performances on different linguistic versions of the test were compared (amongst other measures), it was shown how the respective L scores on each test version varied to a great extent, whereas the ML scores remained relatively stable (El Euch, 2010, p. 28), which, once again, revealed the importance to clearly identify which levels of analysis are being measured in a given task. Furthermore, depending on the perspective adopted by the researcher, it may be justifiable to interpret the L score obtained in this test as an initial stage of awareness, such as discussed by Leow (2000), who distinguishes "high" from "low" awareness, depending on the participant's degree of rule explicitation (Leow, 2000, p. 560). As indicated in the conceptual framework, other researchers have used the notion of "epilinguistic" to refer to such initial stages of awareness, thus pointing to a qualitative rather than a quantitative difference to what can rightfully be taken as metalinguistic knowledge (e.g., Gombert, 2000; Pinto et al., 1999). In sum, a distinct analysis of L and ML scores was thought to be particularly useful for the discussion of the results to allow for interpretations adopting a broader definition of metalinguistic awareness.

It is important to note that half the items in section *(I)* do not allow for a clear identification of the sameness or difference expressed by the given pair of sentences. Thus, only four out of those eight items can be assessed for a linguistic (L) level of analysis. Similarly, only one of the three kinds of figurative language presented in section *(III)*, namely the advertising slogans, can be analyzed with respect to a linguistic dimension, which in this case, is reflected in the comprehension of the slogan's literal meaning. Finally, we included only those items where both linguistic and metalinguistic dimensions of analysis could be assessed. In the first section two of the four items including a linguistic dimension were part of the same subsection, namely "qualitative relations", from which only one item was selected. Accordingly, only one of the two advertising slogans was chosen for this abridged version of the test. More specifically, we discarded the one which, apart from analyzing overt and covert meanings of the given expression, also necessitates cultural background knowledge in order to grasp the underlying humoristic aspect of the slogan. The advertising slogan we discarded namely made reference to a famous quote from Hamlet, which, even though a literary classic, might not necessarily be known to all the participants, which, in

turn, would lead to results indicating a lack of metalinguistic analysis for those who do not grasp the underlying humor of "Être ou ne pas être là, là n'est pas la question" to advertise an answering machine. By contrast, "N'eau fatigue, n'eau stress" to advertise a brand of mineral water, does not include a cultural or literary reference and would thus not penalize students who have different degrees of literary knowledge.

Finally, three (3) items were retained in the *Comprehension* section, one (1) advertising slogan from the *Figurative language* section, and the complete text presented in the *Acceptability* section. However, three (3) of the fifteen (15) errors included in the text were not considered for analysis in our study (see justification below), which left us with twelve (12) items in this last section.

4.3.4.3 Validity and reliability of the metalinguistic ability test

As participants were French native speakers, the French version of the MAT-3, the "Test d'habiletés métalinguistiques" (THAM) (Pinto & El Euch, 2015), was used.[41] Even if the THAM-3 was subsequently validated with university students in the province of Québec, we ran a pilot test to verify whether cégep students were able to reach levels of MLA that met the demands of the test. A shortened version of the THAM-3 (4 items of the *Comprehensio*n section and the complete *Acceptability* section[42]) was administered to five first-year students at *Cégep Beauce-Appalaches* in April 2013, in order to verify whether the tool was appropriate for the learner type and age group in question. The students who accepted to participate in this test run were expected to follow the instructions and do their best to justify their answers. As anticipated, the five students who were tested obtained considerably higher ML scores for *(I) Comprehension* than for *(II) Acceptability*, which is probably due to the different kinds of metalinguistic knowledge required for the justification part of each of the sections. Whereas the former calls for a rather metasemantic analysis, the latter requires the application of more grammar-based terminology, given the metasyntactic nature of the task. However, most of the errors in the *Acceptability* section were detected and well corrected by all of the participants, which suggested sufficient

---

[41] The authors granted us permission to use this test, since it had not yet been published at the time.
[42] At the time of the pilot test, we had not planned to include items on figurative language, which is why no items from this section were part of the test run.

linguistic knowledge to perform the task. Yet, when analyzing the answers in this section, three items proved to be difficult to rate according to the coding procedures provided by the authors. Two of these cases were due to inconsistencies between the error as it appeared in the text and in the coding procedures, respectively, whereas the third was considered ambiguous with respect to the possible interpretations that may have served as an appropriate answer. Consequently, the three errors in question were not included in our final version of the test. Apart from these minor observations regarding the application of the suggested coding procedures, the overall results of the pilot study showed that the test could account for different levels of L and ML analysis, respectively.

In sum, we concluded that a shortened version of the THAM-3 would allow measurement of MLA as a factor distinct from language proficiency. Participants with higher levels of MLA were expected to make more flexible use of their other languages and especially of the typologically related L2, English, in order to make sense of new German words, as assessed by a word translation task (see next section). This hypothesis was mainly based on personal observations of learners' strategies in classroom situations where they are confronted with unknown vocabulary. More specifically, those who try to find analogies with helpful L2 material might not necessarily be those who are very good in English. According to Jessner (2008a), the tendency to make use of supporter languages in L3 learning is one of the constituents of "crosslinguistic awareness", which, in turn reflects the "dynamic interplay between crosslinguistic interaction and metalinguistic awareness" (Jessner, 2008a, p. 279). In Jessner's perspective, however, there is a distinction between this "tacit awareness shown by the use of cognates" and an "explicit awareness in the case of switches introduced by metalanguage" both of which constitute what she calls *crosslinguistic awareness* (see above). Even though we were not interested in language switches observed in L3 production, Jessner's view on crosslinguistic interaction in L3 learners presupposes that this *tacit* awareness leading students to manipulate their own language use, yet without naming the underlying mechanisms explicitly (such as reflected by the L scores of THAM-3), may suffice to make efficient use of supporter languages. We thus considered both the L and ML scores obtained on the THAM-3 in order to point to the contribution that different qualitative (L vs. ML) and quantitative (low vs. high) levels of

awareness would make on the degree to which learners make use of their English to make sense of unknown words in German.

Finally, in line with Herdina and Jessner's (2002) notion of multilingual proficiency (see section 2.1.2.2 on the *Dynamic Model of Multilingualism*), distinct levels of MLA were also expected to become apparent in the verbalizations of the learners' strategies when trying to infer meaning of new L3 vocabulary. Therefore, in order to address the relationship between the learners' metalinguistic abilities and their active use of linguistic resources from different languages from a dynamic perspective, we also coded the TAPs for levels of metalinguistic analysis (see section 4.5 on *Data analysis*).

4.3.5 English proficiency test

4.3.5.1 Rationale for the English proficiency test

Given that L2 proficiency has been proven to be a crucial factor for learners to resort to that language when learning a given L3 (see Chapters 2 and 3), participants' L2 proficiency was assessed. In other studies in which source language proficiency has been identified as an explanatory factor for language transfer, researchers have mostly classified students according to years of formal instruction (Lindqvist, 2010; Odlin & Jarvis, 2004; Williams & Hammarberg, 1998) or according to results from standardized English proficiency tests such as matriculation exams (Ringbom, 2001, 2007) or from non-standardized tests designed solely for the purpose of classification within the respective study (Peyer et al., 2010; Tremblay, 2006). In the particular context of our study, an extensive measure of all the linguistic skills was not judged to be necessary, since our goal was to identify instances of positive transfer from English (L2) to German (L3), yet only with respect to the comprehension of lexical items. Hence, if L2 proficiency plays a chief role for positive transfer in L3 acquisition, it should mostly be English vocabulary knowledge, rather than, for instance, listening, reading or writing skills that would lead our participants to make correct inferences from the target material in German. However, aiming at isolating the proficiency factor as a predictor of positive transfer, we did not believe it sufficient to assess vocabulary comprehension exclusively. Rather, our aim was to assess the participants' overall level of grammatical and structural knowledge, with a focus on

vocabulary comprehension and use in English. This particular context thus justified the use of the *Michigan Test of English Language Proficiency* (MTELP, Corrigan et al., 1979). The MTELP was developed for non-native adult speakers of English to assess their level of language proficiency in an academic setting. It was initially conceived to measure the ability of students who enroll in an English-speaking university to pursue their studies in their L2, or Lx. Whereas it is actually a part of a larger test battery that includes a listening test, an essay writing task and an oral production test, the MTELP only consists of three sections: grammar, vocabulary knowledge and reading comprehension. It is mostly used as an admission test for non-native speakers of English in different American universities.

4.3.5.2 Description of the English proficiency test

The MTELP (see Appendix F for examples) is divided into three sections, each of which consists of an introductory paragraph with written instructions, followed by one or two example questions (or a paragraph in the reading comprehension part), which are then followed by a set of multiple choice questions. The grammar and the vocabulary part both include forty (40) test items, whereas the reading comprehension part consists of four short reading passages, followed by five (5) multiple choice questions each. Hence, the total test score adds up to one hundred (100). In the grammar part, each test item represents a short dialogue, where one person reacts to another. In the second part of each of these exchanges, one word has been omitted, which the test candidate must complete by choosing one out of four optional answers. These fillers represent different kinds of grammatical structures such as, verb tense, aspect, prepositions, modals, etc. According to the authors, this technique supposedly "favor[s] grammatical production" instead of merely assessing the ability to recognize the correct structure (Corrigan et al., 1979, p. 2). Moreover, some of the items necessitate the specific context of the dialogue to be correctly answered, since parsing alone still leaves a possible choice between two options in those cases. In the vocabulary section, part of the items present the examinee with a choice between four possible synonyms to an underlined word in a given sentence, while others are fill-in-the blank sentences, where a single word has been omitted. This time, a choice has to be made between different vocabulary items of the same word class. The majority of the targeted lexical items are verbs, nouns and adjectives. Finally, the reading comprehension section also presents test

takers with four answer choices for each item. The latter are related to one of the four reading passages, each of which counts between 100 and 350 words and covers various text types such as literary criticism or anecdotes as well as different topics such as, for instance, history or physical science (p. 3). As posited by the authors, the questions have been designed as to make the reading of the respective passages indispensable, even if the subject matter were known (p. 3).

4.3.5.3 Validity and reliability of the English proficiency test

Given that the MTELP had already been widely used as a proficiency measure for young adults who speak English as a second or foreign language, we did not have to revalidate the tool with respect to internal validity. However, since it was designed for 1st year university students in the United States, we had to verify whether *cégep* students in Québec were advanced enough in English (L2) to attain respectable scores on this test, which would further make it possible to differentiate proficiency levels in our sample. The R form of the MTELP had been tested one year prior to the data collection period on four students who shared the characteristics of the targeted population. The raw scores attained by these participants on a total of 100 points ranged from 51 to 76, which, according to the equation table for the respective forms, represented equated scores between 60 and 85 (Corrigan et al., 1979, p. 22). Although the participants of the pilot test had received between 8 and 10 years of English instruction, the actual number of hours of ESL instruction possibly varied depending on the particular orientations of the elementary and secondary schools they had attended. It was thus expected that part of our participants would be able to attain scores superior to 80 %, while others might not reach the 60 % threshold, which was thus confirmed, regardless of the small sample size of the pilot study. Despite the fact that the MTELP recommends that only students with scores of 80 and above be admitted to undergraduate programs (p. 12), the test still revealed itself applicable for our target group, given that our aim was not to send them to an Anglophone university, but to discriminate proficiency levels. As specified in Chapter 5, the raw scores of our research sample (N = 66) ranged from 28 to 99 points, which confirmed the appropriateness of this test to evaluate English language proficiency in the target population.

4.4   Data collection procedures

Upon reception of the official letter of approval from the research ethics board of Laval University (see Appendix G), the above-mentioned test materials were administered to the participants in three separate meetings ranging from fifty (50) to ninety (90) minutes each. The first of these sessions took place in the language laboratory of each institution, given that the word translation task and the accompanying TAPs required the use of computers. Since this first testing session took place in a regular class hour, a period of 50 minutes, the background questionnaire was administered immediately after completion of the translation task, in order to optimize time and space. The other two sessions were held in regular classrooms where students could be seated individually. Different dates were set for each of the sessions in the respective institutions, given that the schedules varied from one school to another. However, the order of administrations was, in most cases, identical for each group.

Table 3. *Overview of testing sessions*

Institution	Translation Task & Background Questionnaire	THAM-3	MTELP
Ahuntsic College			
Group I	23-sept, 16.00 – 16.50	23-oct, 14.00 – 15.30	30-oct, 12.00 – 13.00
Group II	24-sept, 15.00 – 15.50	24-oct, 11.00 – 13.30	11-nov, 12.30 – 13.30
Cégep Limoilou			
Group I	29-sept, 9.00 – 9.50	5-nov – 24-nov*	5-nov, 12.00 – 13.00
Group II	29-sept, 13.00 – 13.50		12-nov, 12.00 – 13.00
Group III	29-sept, 17.05 – 17.55		18-nov, 11.00 – 12.00
Cégep de Sainte-Foy			
Group I	20-oct, 10.00 – 10.50	24-nov – 1-dec*	26-nov, 14.00 – 15.00
Group II	21-oct, 11.00 – 11.50		26-nov, 14.00 – 15.00
Cégep Beauce-Appalaches	27-oct, 12.40 – 13.35	20-nov – 27-nov*	20-nov, 14.30 – 15.30
Group I			27-nov, 14.30 – 15.30
Cégep Lévis-Lauzon			
Group I	21-nov, 15.00 – 16.50	19-nov – 28-nov*	19-nov/26-nov
Group II	26-nov, 16.00 – 16.50		19-nov/26-nov
Group III	28-nov, 15.00 – 16.50		19-nov/26-nov

* Given the difficulty of finding a date and time where everybody was available for a second meeting outside the class hours, we arranged separate testing sessions with smaller groups to administer the THAM-3 throughout the listed periods.

4.4.1 Translation task and background questionnaire

With the exception of the students from *Cégep Lévis-Lauzon* who had been recruited later into the term, we had arranged with all of the other institutions that the translation task would be administered in a regular German class, taking place in a language laboratory, generally used for speaking and listening practices in German. All students would thus participate in this testing session whereas the researcher would only collect the recordings

of those who had signed a consent form to participate in the rest of the study. For the other students, this translation task would simply represent a certain amount of exposure to German written and spoken language from which they would possibly pick up some new vocabulary in the target language. After the translation task was completed by all of the students of each group, only the participants of our study were then asked to fill out the background questionnaire, while the rest of the group was free to leave the language laboratory.

4.4.1.1 Ahuntsic College

We met the first group of students from *Ahuntsic College* on September 23 between 16.00 and 16.50, which was their regular class hour in the language laboratory. Out of the 19 students who were present that day, 14 had accepted to participate in the rest of our study. Given that this first test took place in their regular class hours, the entire group took the test, but only the data from those participants who had signed the consent forms were collected to be analyzed as part of the sample. Each student was placed in front of a computer in a language laboratory. First of all, a few technical details had to be set up. The participants were invited to verify the sound settings on their computers and to try recording their voice, which they were then asked to listen to, in order to assure that their verbal protocols would be recorded properly. Then, the researcher reminded the participants of the nature of the test, of which they had previously been informed in the consent forms. At this point, the researcher explained in French how this translation task was going to work:

> « Je vais vous demander de traduire des mots de l'allemand vers le français. C'est normal que la plupart des mots vous soient inconnus. Essayez d'en faire sens du mieux que vous pouvez. Chaque mot va apparaître pendant 30 secondes sur l'écran et vous allez aussi l'entendre et le voir dans le contexte d'une phrase. Pendant que vous êtes en train de chercher la bonne traduction, SVP, pensez tout haut, c'est-à-dire prononcez à voix haute tout ce qui vous passe par la tête! Peut-être que vous avez déjà appris l'un ou l'autre de ces mots ou que ça vous rappelle quelque chose d'autre. Mais il se peut aussi que ça ne vous dise rien et que vous n'ayez aucune idée de ce que ça veut dire. Dans les deux cas, dites tout ce qui vous vient à l'esprit

pour trouver la bonne traduction. Vos réponses seront enregistrées automatiquement. »

[You will be asked to translate a number of words from German into French. It is expected that most of these words will be unknown to you. Try to make sense of them as best as you can. Every word will be shown on the screen for 30 seconds, you will hear a pronunciation of the word and see it in a sentence context. As you are trying to find an appropriate translation, please speak out loud all the thoughts that come to your mind! You may already have come across one or the other of these words. However, it is also possible that they do not relate to anything you know and that you have no idea what they mean. In any case, state everything that crosses your mind while you are searching for a translation. Your answers will be recorded automatically.]

Then, the researcher led the participants to the test interface, by showing, step by step what they had to do to start the test, on a large screen displayed in the front of the room. First, they had to copy the test file, including the sound files of the test items, the instructions and the script to be run in DmDX, from the directory on each of their screens. Then, they were asked to rename this file using their own first and last names, in order for us to be able to identify their recordings after the test was done. Subsequently, they could launch DmDX, which had previously been installed and tested on all the computers in the laboratory. The first interface to appear was a page with written instructions (see Appendix D), after which each participant was free to start the test by pressing the space bar. The 42 test items then appeared in random order and each was displayed for 30 seconds. As the participants started thinking aloud while they were trying to translate the items, the researcher went through the classroom listening to what was being said. On a number of occasions a participant would simply translate a word without elaborating on it any further, in which case the researcher would ask him or her, how they had come to this particular answer. This would lead the participants to specify whether they had already learned or heard the given word, whether their answer was based on an association with a similar word from another language or whether they were simply guessing the translation. Of course, this intervention may seem arbitrary, given that it was possible that not all of these cases were noticed by the researcher. However, the analysis of the TAPs revealed that all the translations that were

given without an elaboration on the source of knowledge included markers of uncertainty with respect to the translation choice, thus suggesting that the correct translation was not previously known (see section 5.2.2). When the translation task was finished, the researcher handed out the background questionnaires to all those who were to participate in the rest of the study, while the rest of the group was then free to leave. As the participants filled out the questionnaires, the researcher collected the data from the computers by means of USB-sticks.

We met the second group of German learners (N = 5) from *Ahuntsic College* the following day, September 24, at 15.00. Again, the test took place during their regular German class in the same language laboratory in which the first group had taken the test the day before. Five out of seven students from this group were present. The main procedural steps were identical to those of the previous session. However, in the oral instructions, we specified that participants were invited to elaborate on how they knew certain translations, when they seemed confident about their answers without indicating whether or not they had learned the given word before. The following instruction was inserted before the final statement about the automatic recording:

> « Si vous connaissez déjà la bonne réponse, indiquez comment vous le savez, vous l'avez peut-être déjà appris en classe ou ailleurs. »
>
> [If you already know the correct answer, indicate how you know this. You may already have learned it in class or elsewhere.]

Again, the researcher walked around the classroom to listen to their verbalizations after the participants had started the test. It happened only once that the researcher intervened with a specification question after a participant had translated a word without elaborating upon the origin of his or her knowledge. As soon as they had finished working on the translation task, the participants were asked to fill out the background questionnaire. During that time, the researcher collected their oral data by means of USB-sticks.

4.4.1.2 Cégep Limoilou

On September 29, we met the first group at *Cégep Limoilou* at 9.00 in the language laboratory. The entire group (N = 20), 13 of whom had accepted to participate in our study,

followed our instructions attentively and seemed keen on finding the right translations for the German words that they were asked to reflect upon. Again, the procedural steps were mostly identical to the previous testing sessions at *Ahuntsic College*, only that the explanations prior to running the test on these computers were slightly different due to the technical requirements of this language laboratory. More precisely, no sound tests were necessary given that we had tested each station prior to the students' arrival in class. Hence, we only had to ask the participants to retrieve the test material before explaining the nature of the test, including the specification on previously acquired vocabulary (see above). Again, participants were then free to start running the test when they were ready, after having read the written instructions on the screen. The session went smoothly and only a few oral interventions on the part of the researcher were necessary to have some participants elaborate on the origin of their knowledge, when they were convinced about an answer without indicating how they knew this. When the task was over, the participants were asked to save their data and drop the audio files into a specific folder on the server, which was specified to them by the researcher. This enabled us to copy the entire data set from one folder, instead of collecting the recordings manually from each computer. Again, the last step of the meeting involved filling out the background questionnaires.

We met the two other groups from *Cégep Limoilou* later the same day, in the same language laboratory. Out of the 22 students from the second group, 15 had accepted to take part in the project. However, one of them was absent from school that day and one other participant had to leave the classroom due to anxiety issues, once the students around her had launched the test. Therefore, we retained 13 participants from this second testing session, which was held from 13.00 until 13.50. Except for the one student who left the laboratory after approximately 10 minutes[43], the session went smoothly, and the questionnaires were handed out after the task, as in the previous group.

As for the last group, who we were supposed to meet at 17.00 in the language laboratory, we had to change our plans slightly, given that not even half the group had arrived by 17.05. We decided to administer the background questionnaire to those who were already

---

[43] This participant was not included in the sample.

present and handed it out as the remaining students gradually arrived. By 17.15, all the students had arrived and kept on completing the questionnaire until 17.25. At that time we started giving the oral instructions and everything went smoothly from there, even if the group was bound to stay an extra five minutes to finish the test, which was completed by 17.55.

4.4.1.3 Cégep de Sainte-Foy

In this institution, the translation task was administered in week 8 of the German course, on October 20 and 21. This was not only for logistical reasons but also because these students received 3 hours of German instruction per week, as opposed to 6 hours per week at *Ahuntsic* and 4 hours per week at *Limoilou*. Therefore, by holding the testing session three weeks later than at *Limoilou*, we made sure that all participants had received approximately the same amount of German instruction, when the translation task was administered. There were 7 out of 26 students from the first group and 9 out of 21 from the second group who gave their consent to participate in the rest of the study. The testing sessions went smoothly in both groups. They were held in the same language laboratory and no technical problems arose during the test. The instructions given by the researcher before and throughout the test were identical to those given in the previous sessions, including the specification concerning potential cases of prior knowledge (see section 4.4.1.1). When the students had finished the test, they were asked to identify their audio files and to copy them on USB sticks handed out by the researcher. In the meantime, students who had agreed to participate in the study were asked to fill out the background questionnaire. Each session, including both the translation task and the questionnaire, took up the entire 50 minutes of the class hour. Unfortunately, two recordings from this testing session were lost due to technical problems related to computer settings. As a consequence, the two participants in question had to be excluded from the sample.

4.4.1.4 Cégep Beauce-Appalaches

At *Beauce-Appalaches*, we met the students on October 27, one week after the previous session at Sainte-Foy. There were 23 students who had all accepted to participate in our study. However, one student dropped out of the program later in the term and another

decided to end her participation, because it was too time consuming for her. When the class started at 12:40, all the students were present and the researcher gave the same explanations as in the previous groups. An initial recording problem led the researcher to hand out the questionnaires before the translation task, while the technical support service helped fixing the problem. At 13:10, the test could be run with the appropriate sound settings. Except for one participant, the collected recordings were complete and usable. Thus, the final number of participants from this group was 20.

4.4.1.5 Cégep de Lévis-Lauzon

This last college was contacted in the last week of October in order to recruit additional students, given that the coding of the questionnaires had led us to exclude a large number of participants with language backgrounds that deviated too much from the targeted profile (see *Participants* section). Nine additional participants were recruited from *Lévis-Lauzon*. On November 21, 26 and 28, we arranged three meetings in small groups of three participants who were seated at the extreme ends of the classroom, in order to leave as much space as possible between them. The researcher brought laptops on which the program had previously been installed to insure proper functioning. The instructions were given orally, as in the other groups. However, the researcher specified that a certain number of test items would probably already be known to the participants, given they had received considerably more hours of German instruction than the participants from the other schools, namely between 36 and 42 hours, as opposed to a maximum of 24 hours in the other institutions. The researcher simply underlined the fact that it was important to state how they had learned the respective item if they already knew a translation. Even if the technical surrounding was not ideal, the students managed to concentrate on their task and to follow the instructions. Unfortunately, one of the nine recordings that were collected immediately after the testing session was not usable due to the poor quality of the recording.

4.4.2 Metalinguistic ability test – Test d'habiletés métalinguistiques (THAM-3)

In a second meeting that was scheduled one to four weeks after the translation task, the researcher administered the abridged French version of the Metalinguistic Ability Test, the THAM-3, for which two fifty-minute periods had been reserved for each group in all five

institutions. However, given that this meeting took place outside of the regular class hours, it was not possible to meet the same groups of participants at the same time. Table 3 (see section 4.4) summarizes the periods of time during which the different meetings were held in each institution. However, the conditions in which this test was taken were nearly identical for all participants. All the testing sessions took place in an empty classroom and consisted of a silent pen and paper test. Each participant was seated alone at one table and only the total number of participants varied from one testing session to the other.

Following the authors' recommendations, participants were seated at individual tables, in order not to be disturbed by their peer students, given that the test necessitated sustained concentration (Pinto & El Euch, 2015). The three sections were administered in the following order:

1. *Comprehension* (3 items, 30 minutes)
2. *Figurative Language* (1 item, 10 minutes)
3. *Acceptability* (40 minutes)

The researcher read out loud the instructions and the example for the first item of the comprehension section, both of which were also printed on the test sheets (see Appendix B), so that participants could validate their answers when in doubt. Furthermore, the examiner specified the time allocated for each item, namely ten minutes, while indicating that time did not constitute an evaluation criterion, but that a limitation was necessary to make sure that all participants would be able to finish all sections. Once the first ten minutes had passed, the participants were reminded to turn to the following item in case some of them were still working on the first one. After approximately thirty minutes, they were asked to turn to the one item of the section on figurative language, namely an advertising slogan. The researcher read out loud the instructions and reminded the participants that they disposed of ten minutes to work on this item. Even if sections 1 and 2 did not exceed the first fifty minute-period, which allowed participants to take a short break before the last part of the test, all of them stayed in the class to complete the *Acceptability* section immediately. The participants were told that they disposed of a total of forty-five minutes to work on this task, in which they had to detect and correct errors in a short text and justify their answers for each error (see above).

4.4.3 Michigan Test of English Language Proficiency (MTELP)

The last meeting was devoted to the Michigan Test of English Language Proficiency (MTELP). Similarly to the THAM-3, separate meetings were arranged for different groups in the five institutions. Again, this was one to four weeks after the translation task (refer to Table 3 in section 4.4 for the exact dates). Apart from a few exceptions[44], the MTELP was the last test on the schedule. As for the THAM-3, the conditions under which the MTELP was administered were identical for each group in all the institutions, with the only difference being the size of the group. Again, the students were seated individually and asked to work on the test silently.

As suggested by the authors of the test, a total of ninety minutes was allocated to answer all the questions in the three sections. The MTELP was administered as a whole, in order to give participants who took less time answering the grammar and vocabulary questions the chance to continue directly with the reading section, which would allow them to leave as soon as they had finished. According to observations from the test run, it seemed probable that a number of students would finish within sixty to seventy minutes, whereas others were expected to take the whole ninety minutes to work through the test. Surprisingly, those who took over 80 minutes to complete the test were the exception while the majority of the participants needed approximately 60 minutes for the whole test. Some even took less time. Similar to the procedure for the previous materials, the test taker read out loud the instructions of the first section, explained the following sections and reminded the participants of the total time available to answer the questions.

4.5 Data analysis

The goal of our study was to point to the most influential factors that might lead beginning learners of German (L3) to understand unknown vocabulary items in German by recognizing their similarity to cognate words from English in a translation task. More

---

[44] There was a total of 8 participants who were not able to come to any of the suggested testing sessions for the THAM-3, scheduled before those for the MTELP. These participants were met individually at a subsequent date to take the THAM-3. This explains the periods of time indicated for the THAM-3 (see Table 3), which go beyond the date at which the MTELP was administered in four of the five institutions.

precisely, we tried to determine which of the main three explanatory variables identified above, namely proficiency in English (L2), exposure to English (L2), and metalinguistic awareness, was the strongest predictor of positive transfer from English, as measured by a translation task, the results of which were examined by secondary data from think-aloud protocols (TAPs). Along with these central factors that were expected to influence positive transfer, we included the level of interest in the German language and culture as two further potentially influential factors, in order to add a perspective on the affective dimension of transfer phenomena.

4.5.1 Coding procedures

The table below provides an overview of the different types of ratings that were applied to determine partial and total scores for each variable, as measured by the respective instruments. Whereas the coding procedures of the MTELP and of the questionnaire are thus clearly outlined, the respective analyses of the think-aloud protocols (TAPs) and the THAM will necessitate some further elaboration. For instance, the TAPs served as a basis of analysis for two distinct variables. In the first place, it was used to determine 'Positive transfer from English'. However, we also analyzed the verbalizations for 'levels of metalinguistic analysis', as will be explained in more detail in the following subsection (4.5.1.1). Still, the THAM-3 remained the primary measure of 'metalinguistic awareness'. The coding procedure of this instrument will be briefly summarized under 4.5.1.2 and discussed with respect to the relative levels of MLA determined in the TAPs.

Table 4. *Coding procedures (overview)*

Variable	*Instrument*	*Type of rating*	*Partial scores*	*Total score*
Positive transfer	TAP	Incorrect/absence of translation (or prior knowledge) = 0 Correct translation = 1	Maximal score per item = 1 (28 items)	28 (max)
Metalinguistic awareness	TAP	Absence/prior knowledge = 0 Intuition = 1 Noticing = 2 Meta-Awareness = 3 Underlying Rule = 4	Maximal score per item = 4 (28 items)	112 (max)
	THAM	Linguistic answers: - incorrect = 0 - correct = 1 Metalinguistic answers: - no analysis = 0 - partial analysis = 1 - in-depth analysis = 2	L score = 16 (16 items) ML score = 32 (16 items)	48 (max)
English proficiency	MTELP	Multiple choice test - incorrect = 0 - correct = 1	Grammar = 40 Vocabulary = 40 Reading = 20	100 (max)
Exposure to English: *Frequency of use*	Questionnaire (Q21)	Never = 0 Several times a year = 1 Several times a month = 2 Several times a week = 3 Daily = 4	Maximal score per item = 4 (10 activities)	40 (max)
Exposure to English: *Basic exposure*	Questionnaire (Q8)	Never = 0 One week or more = 1 Four weeks or more = 2 Three months or more = 3 One year or more = 4	Maximal score per item = 4 (4 environments)	16 (max)
Interest in German language Interest in German culture	Questionnaire (Q19) Questionnaire (Q20)	"très faible" = 1 "faible" = 2 "moyen" = 3 "fort" = 4 "très fort" = 5	Maximal score per question = 5	5 (max)

4.5.1.1 Coding of the think-aloud protocols (TAPs)

First of all, 'Positive transfer from English' was determined by calculating the number of correct translations for the 28 test items whose meaning could be inferred on the basis of an English cognate. However, when the analysis of the verbal protocol revealed that a correct answer was based on the participant's prior knowledge of the test German item, without associating the item to its English translation equivalent, the answer in question was excluded. Correct translations were coded as instances of positive transfer when one of the following conditions was met: (1) the association to English was explicitly mentioned, or (2) the association to English was implicit, when no other source of knowledge was mentioned. Since transfer rates were calculated for the 28 items for which English translation equivalents could serve as a basis for positive transfer, it was possible to obtain a maximum score of 28. Total transfer rates were normalized for each participant after exclusion of answers based on prior knowledge.

Only afterwards were the TAPs coded for depth of analysis, in order to complement the primary measure of MLA and thus to serve as a means of data triangulation. Whereas the THAM-3 measured the participants' more general metalinguistic abilities, as evidenced in their manipulations of different structures in their native language, French, the TAPs provided an insight into their capacity of applying such analytic skills to a multilingual context. The verbalizations therefore provided evidence for "crosslinguistic awareness" (Jessner, 2008a), specifically in cases where analyzed knowledge about the structural features of different languages was actively used to justify inferencing strategies such as the recognition of cognates, the identification of syntactic functions or phonological patterns of a given target item.

The coding procedures used for the TAPs collected in the present study were based on the few studies in SLA that have used this instrument to determine levels of metalinguistic processing during different problem-solving tasks in a second or additional language (Hama & Leow, 2010; Leow, 1997; Rosa & Leow, 2004; Rosa & O'Neill, 1999; Sachs & Suh, 2007). The overall aim of most of these studies was to investigate whether reported levels of awareness had a differential effect on the recognition and/or the accurate use of

specific TL structures. Except for Hama and Leow (2010) who rather addressed the possibility of learning without awareness, the other studies used pre-test/post-test designs targeting specific grammatical features[45]. All used concurrent think-aloud protocols (TAPs) to assess levels of awareness, based on Schmidt's (1990, 1994, 2010) *noticing hypothesis*, which states that conscious attention at the level of 'noticing' is necessary for learning to take place. This initial stage of awareness has been operationalized as a "cognitive operation that takes place both during and immediately after exposure to the input that is available for self-report" (Schmidt, 1990, p. 132). It is typically opposed to the level of 'understanding', which is most often associated with the explicit formulation of rules, and thus represents a deeper level of awareness (see section 2.3.1.2 on the notion of awareness). From the studies above, some have differentiated between these two levels of awareness for the coding procedures of their TAPs. Rosa and O'Neill (1999), for instance, identified any "verbal reference to the target structure without any mention of rules" (p. 529) as noticing, whereas understanding was "operationalized as an explicit formulation of the rule underlying the target structure" (p. 530). The studies by Rosa and Leow (2004) as well as Hama and Leow (2010) distinguished the same two levels of awareness, yet including a third category named "no report" for all instances where "participants failed to provide evidence of being aware of the target form" (Rosa & Leow, 2004, p. 280).

As briefly outlined when elaborating on the notion of awareness in the conceptual framework (see section 2.3.1.2), Leow and his colleagues (2011) have suggested a finer-grained analysis by identifying an intermediate stage of awareness, which goes beyond mere noticing, yet does not entail explicit rule formulation. Just such an elaboration is evidenced in the empirical investigations of Leow (1997) as well as Sachs and Suh (2007)[46], who differentiated between the same three levels of awareness. Beyond the first category labeled Cognitive Change (CC) at the level of *noticing*, the intermediate stage labeled Meta-Awareness (MA) entails a report of the learning experience in which participants explicitly state that they see a certain pattern emerge (see Leow, 1997, p. 478).

---

[45] Some looked at stem changes in preterite verbs forms in Spanish (Leow, 1997), others at conditional sentences in Spanish (Rosa & Leow, 2004; Rose & O'Neill, 1999) or backshifting in English (Sachs & Suh, 2007).

[46] Sachs and Suh (2007) based their coding procedures on Leow's (1997) and operationalized the three levels accordingly (Sachs & Suh, 2007, pp. 210-211).

Hence, even if this intermediate stage does not include the metalinguisitc description of an underlying rule, as represented by the third category labeled Morphological Rule (MR), verbalizations coded at the level of MA are taken to reflect the level of *understanding* (see Table 5).

Table 5. *Coding of TAPs (adapted from Leow, 1997)*

Label	Level of awareness	Operationalization
CC Cognitive Change	Noticing	Some behavioral or cognitive change due to the experience
MA Meta-Awareness	Understanding	Report of being aware of the experience; recognition of some pattern
MR Morphological Rule	Understanding	Some form of metalinguistic description of the underlying rule

Based on a number of examples from their sample, verbalizations were qualified as reflecting Meta-Awareness (MA) when they included justifications based on the recognition of a pattern, yet without reference to grammatical rules (Leow, 1997, p. 480). After the initial coding procedure, the researcher further split the sample into a meta-awareness and a non-meta-awareness group, based on this second criterion. The results revealed that meta-awareness, whether this included explicit rule formulation or not, was strongly associated with recognition of the targeted forms and also with the performance on a production task involving these forms in the post-test (p. 484). As a consequence, what these observations indicate is that a finer-grained analysis of awareness might be the key to better grasp its role for the learning process of non-native languages.

The coding procedures that we adopted for the qualitative analysis of the TAPs in the present study were mainly based on these three levels. However, we retained the label 'noticing' (N) instead of 'cognitive change', which seemed more appropriate to the specific task type, given that it entailed 'noticing of formal resemblance to another word', rather than a change in verbal behavior. Moreover, we renamed the last category 'underlying rule' (UR) since the word translation task did not imply the manipulation of morphological patterns but rather of sound-spelling correspondences or of certain semantic features evoked by the lexical items themselves or the co-occurring words in the sentence. Finally,

given the particularity of our task, which involved translating words based on perceived similarities, we added a category labeled 'intuition' (I), which referred to random clang associations, i.e., where similar-sounding items were associated without any further reflection.[47] This specific label was used with reference to developmental psychologists such as Tunmer and Herriman (1984) who used the term "intuitive awareness" (p. 19) to refer to epilinguistic processes which do not involve the subject's conscious control over their manipulation of language (e.g., Gombert, 2000, as discussed under 2.3.1.1).

As shown in Table 6 below, the four categories thus indicating minimal to maximal awareness in ascending order (I, N, MA, UR), were rated 1 to 4, respectively. As illustrated by the examples discussed in Chapter 5, codings in these four categories did not depend on whether or not the given item was correctly translated or whether it had been transferred from English. In other words, each verbalization was coded independently for 'transfer from English' and 'level of metalinguistic analysis', respectively. However, two additional categories were added to cover, on the one hand, cases where the item was already known, labeled 'prior knowledge', and, on the other, those where no translation attempt was made, or else where such an attempt did not reflect any kind of reasoning, which we labeled 'absence'. Both these categories were rated 0. Table 6 illustrates the six categories according to which the TAPs were coded (see Appendix E for a more detailed version).

---

[47] As will be exemplified in Chapter 5, it was possible to translate a word correctly based on such an intuitive association, which was coded as 'intuition' rather than 'noticing' when no explicit reference was made to the English cognate but when the answer was clearly based on the formal similarity to the latter.

Table 6. *Coding of the TAPs*

Label	Operationalization	Rating
PK Prior knowledge	- of translation equivalent w/o statement of resemblance	**0**
A Absence	- of translation - of reasoning	**0**
I Intuition	- of correct translation based on English cognate w/o explicit reference - random clang association w/o logical reasoning	**1**
N Noticing	- of formal resemblance to English cognate - of resemblance to another word that could have fit the context	**2**
MA Meta-Awareness	**Report of line of thought** - orthographic and/or phonetic resemblance explicitly stated and taken as justification - reasoning based on internal logic - cross-check with sentence or other learning events - reference to typological proximity/distance	**3**

UR
Underlying rule
	Report of underlying rules and/or related meanings
	- justification based on
	  o explicit formulation of phonetic/orthographic pattern
	  o syntactic/semantic properties of co-occurring or associated lexical items
		4

4.5.1.2 Coding of the THAM-3

As for the primary measure of MLA, test results were first analyzed for each of the dimensions separately in order to obtain two independent test scores. Whereas the linguistic answers were rated according to yes/no answers or accurate/inaccurate error corrections, the metalinguistic answers were analyzed qualitatively following the scoring procedures suggested by the authors of the test. Accordingly, each answer was rated on a 3-point scale from level 0 (no analysis) to 1 (partial analysis) or 2 (thorough analysis). Furthermore, the authors provided a number of examples for each level of analysis for all the test items. Nevertheless, two independent raters were responsible for the scoring of this part of the test, in order to assure non-bias and inter-rater reliability. Finally, the results obtained on the linguistic questions of the test (L score) were added to those obtained on the metalinguistic questions (ML score) to form a combined score. The latter was taken as a global measure of metalinguistic awareness, namely including the lower levels of analysis which have also been called "epilinguistic" (e.g., Gombert, 2000), representing an initial

stage of awareness. More specifically, in line with the observations regarding the coding of the TAPs (see previous section), we consider Leow's distinction between "low" and "high awareness" (Leow, 2000, p. 560) very helpful since insufficient mastery of grammatical terminology, i.e., explicit knowledge of underlying rules, as shown in the justification part, would not necessarily hinder learners from consciously manipulating the language and could thus reveal a certain degree of MLA. Therefore, we adopted a slightly broader definition of metalinguistic awareness than the authors of the THAM-3, which allowed for an inclusion of such initial stages of awareness, in which participants were not necessarily capable of stating underlying rules explicitly. Even if initial stages of awareness, as reflected in the L-score, were included in the overall rating of the test, a higher overall score naturally reflected a larger number of justifications based on more analyzed knowledge, given that for most items, a correct L answer was the basis for a successful justification, as evidenced by the ML score. In other words, the combined score reflects stages of awareness ranging from a rather intuitive manipulation of the native language to an in-depth understanding of its underlying structures, including syntactic and grammatical features as well as semantic and pragmatic functions of words in different contexts.

In sum, the coding procedures of the THAM-3 and those of the TAPs reflect the different levels of MLA in a similar way. Namely, depth of metalinguistic analysis is seen as a continuum ranging from intuition to explicit formulation of underlying patterns. However, the THAM-3 measures the learners' reflexive awareness, in the sense that they reflect on features of their native language that they are not currently using. On the other hand, the type of awareness that is evidenced in the verbalizations during the word translation task corresponds to the application of metalinguistic abilities to a specific learning context in which the participants activate their linguistic resources across languages. In fact, this "crosslinguistic awareness" (Jessner, 2008a) may be seen as a sort of MLA in action. In other words, the successful manipulation of resources from different languages appears to depend on basic metalinguistic abilities that "mak[e] objectification possible" (Jessner, 2008a, p. 279). As a consequence, the more *reflexive* type of awareness, as evidenced by the integrated score of the THAM-3, was expected to correlate with the *applied* awareness exhibited in the participants' verbalizations, i.e., the metalinguistic score of the TAPs.

Moreover, by investigating the respective correlations of these two separate measures of MLA with that of positive lexical transfer, we hope to contribute substantially to a better understanding of the depth of analysis required for this type of crosslinguistic inferencing.

4.5.2 Statistical analyses

After the coding procedures were completed, the data were submitted to a number of statistical analyses. First of all, we ran a Student-t test to determine whether participants significantly resorted to transfer from English. Only then did we proceed to a regression model to determine which of the independent variables could be identified as a predictor of positive transfer. Given that the quantification of the dependent variable, namely 'positive transfer from English', involved a number of successes in a fixed number of trials, a model of binary logistic regressions was used as the main statistical procedure to investigate the relative contribution of the predictive variables. As will be discussed in the last section of Chapter 5, we ran two independent regression models, the first of which featured the THAM-3 as the exclusive measure of MLA, whereas the MLA scores from the TAPs were only added into the equation in the second. This procedure enabled us to determine whether the reflexive dimension of MLA alone, as measured by the THAM-3, would predict the response variable.

5 Results

As previously stated, the goal of the present study was to investigate which of the variables that have been shown to impact transfer phenomena in multilingual contexts would predict positive lexical transfer from English (L2) to German (L3). More specifically, we sought to determine whether MLA would be a stronger predictor of such positive transfer than L2 proficiency and L2 exposure respectively. At the beginning of this chapter, the descriptive test results are reported with respect to the central variables that were targeted by the different instruments administered to our participants throughout the study. Our sample consisted of 66 French-speaking learners of German who had all started to learn English as a second language at primary school, who had received formal instruction in English for an average of ten years (see Table 1) and who were all beginners in German. All participants took all the tests under nearly identical conditions (see previous chapter).

Following a brief overview of the overall test results of the word translation task, the first subsection pertains primarily to the instantiation of positive lexical transfer from English, as evidenced by the verbal protocols that accompanied the translation task. The second subsection provides a detailed account of the participants' levels of MLA, as determined by the partial and total scores obtained on the metalinguistic ability test (THAM-3), on the one hand, and by means of a qualitative analysis of the TAPs, on the other. The third subsection then briefly reports on levels of English (L2) proficiency and exposure, which were assessed by means of the MTELP and of a questionnaire, respectively. Having thus presented the descriptive test results variable by variable, we will present and discuss the regression model that was used to determine what variables predicted positive lexical transfer from English.

5.1 Word translation task

Prior to discussing the test results that are specifically related to the dependent variable, positive lexical transfer from English, we will briefly present the descriptive results of the

translation task on the whole. It was used to collect data on the processing mechanisms of beginning learners of German (L3) who already spoke English (L2) and some Spanish (L3) besides their native language French (L1), when they are faced with new vocabulary in German. As previously stated, the translation task contained twenty-eight (28) test items that could be related to English due to typological similarity. Additionally, we included six (6) items that were Latinate borrowings, which could thus be traced back to either French or Spanish cognates, and eight (8) items that were typologically unrelated to any of the participants' background languages. The following subsection provides a brief overview of the test results.

5.1.1 Overview of the test results

The three tables below illustrate the descriptive results of the translation task for each group of words separately. More precisely, each table provides a summary of the different kinds of answers that were given for the twenty-eight German-English cognates (Table 7), the six Latinate borrowings (Table 8), and the eight remaining items (Table 9) that were typologically unrelated to any of the participants' background languages. At first, answers were divided into correct versus incorrect translations, as well as omissions or instances where no translations were given. Then, the correct translations were further divided into instances of positive transfer and prior knowledge, respectively. For each of these categories, we provided minimum, maximum, and mean scores, as well as standard deviations across the sample (N = 66). As illustrated in Table 7, correct translations of the German-English cognates ranged from 5 to 26 out of 28. In other words, none of the participants attained the maximal possible score for this group of words, yet some came quite close. However, a closer look at the second and third lines of the table reveals that correct translations were partly due to prior knowledge of certain test items. More precisely, the maximum of correct translations that could actually be identified as instances of positive transfer was of 21. On the other hand, at least one participant reported prior knowledge of 7 out of these 28 items that may otherwise have been translated on the basis of an English cognate. Thus, correct answers based on prior knowledge were removed from

each individual observation and the final scores normalized across the sample.[48]

Table 7. *Descriptive Statistics (Translation Task: German-English cognates)*

	Min	Max	Mean	SD
Correct translation (28)	5	26	13.68	3.96
• *Positive transfer (English) (28)*	4	21	11.86	3.47
• *Prior knowledge (28)*	0	7	1.82	1.75
Incorrect translation (28)	0	18	9.45	4.57
No translation / omission (28)	0	17	4.86	3.94

One specification seems crucial with regard to the instances that were identified as based on prior knowledge. Namely, one might have expected the eight (8) students from *Lévis-Lauzon* who were recruited several weeks into the semester and who had already received between 36 and 40 hours of German instruction, as opposed to a range from 20 to 24 hours for the rest of the sample (see section 4.2) to affect the overall test results with respect to prior knowledge. Interestingly, however, an exclusion of these slightly more advanced learners affected this variable only marginally. That is to say, the mean for prior knowledge dropped from 1.82 to 1.64, while the minimum (0), maximum (7) and mode (0) remained the same. Among the eight participants in question, two had no prior knowledge of any of the items, and only one had already learned seven (7) of them. However, this was also true for one of the participants of another group who were, technically speaking, less advanced. In other words, the fact that some of the items were already known did not appear to be directly related to the amount of German instruction that these learners had received. Since participants who were subsequently added to our sample did not differ significantly from the rest of the group with respect to prior knowledge, all were retained for the totality of the observations in our protocol.

Let us now proceed to a brief overview of the other types of words that were included in the translation task, besides the German-English cognates. As illustrated by the maximum and mean scores in Table 8 below, the six Latinate borrowings appeared to be slightly easier for

---

[48] For example, the total number of instances based on positive transfer of a participant who gave seven correct answers based on prior knowledge (and not on transfer), would be calculated on a total of 21 instead of 28.

our participants to translate on the basis of their French or Spanish cognates. Namely, some participants did attain the maximum of possible correct translations on the basis of positive transfer from French or Spanish. Also, the mean score of these correct translations represented nearly 58.8 % (3.53 out of 6), as opposed to 42.4 % (11.86 out of 28) for the German-English cognates (see Table 7 above). However, these percentages are not easily comparable given the considerable difference in totals. Namely, two out of the six Latinate borrowings, namely *Krawatte* [kraˈvatə] and *Tasse* [ˈtasə], were nearly identical to their French counterparts *cravate* and *tasse*, whereas such a high degree of similarity was only the case for *Fisch* [fɪʃ] and *Maus* [ˈmaus] for the German-English cognates. Naturally, these four test items were also the most frequently transferred, as illustrated in Table 10. To be more specific, *Krawatte* and *Tasse* were both related to their French counterparts by 61 out of 66 participants, *Fisch* and *Maus* to their English cognates by 60 and 59 participants, respectively. One could consider these four items as giveaways, since they appeared to be an easy guess for nearly everyone in our sample.[49] From this perspective, it is not unexpected that the mean transfer rate of the Latinate borrowings represented a higher percentage than that of the German-English cognates, given the ratio of "very easy" to "more difficult" items, as far as similarity relations are concerned.

Table 8. *Descriptive Statistics (Translation Task: Latinate borrowings)*

	Min	Max	Mean	SD
Correct translation (6)	1	6	3.67	1.22
• *Positive transfer (French/Spanish)(6)*	1	6	3.53	1.11
• *Prior knowledge (6)*	0	2	.14	.43
Incorrect translation (6)	0	4	1.65	1.14
No translation / omission (6)	0	4	.68	.98

Finally and not surprisingly, what Table 9 reveals about the eight typologically unrelated

---

[49] It was of course anticipated that different degrees of similarity would affect transfer rates. In the case of *Krawatte* and *Tasse*, this was not considered problematic, since the focus of our inquiry was not on the influence of French background vocabulary. As for *Fisch* and *Maus*, even if the pronunciation of these words was identical to their English counterparts *fish* and *mouse*, the orthographical deviation was considered a sufficient condition to raise doubts about the one-to-one relationship between the German test item and its English cognate, and thus to incite participants to some kind of interpretation.

words in our translation task is that they were, on the whole, a lot more difficult to translate for our participants than those items that resembled their translation equivalents in other languages. Similarly to the other types of words presented above, there were also some participants who already knew some of these typologically unrelated words and thus translated them correctly. However, and most interestingly, there were some cases of what we decided to call "transfer of other knowledge" (see Table 9). These referred to correct translations that were based on associations with co-occurring words or structures or other learning events altogether. Thus, the items *Berg*, *Lied*, *Zeitung* and *Luft* were translated correctly by means of some interesting kinds of reasoning. For example, *Berg* was associated to *iceberg* by four participants, who then inferred that the target item had to be something like a hill. Similarly, six participants were familiar with the German word *Luftwaffe* (English: *air force*) or with the name of the airline *Lufthansa* and thus inferred the meaning of Luft (English: *air*) based on intra-lingual associations. In the case of *Lied* and *Zeitung*, the associations were based on co-occurring words and structures in the supporter sentence. Even if they do not strictly belong there, we listed these instances in the column "positive transfer" in Table 10. Apart from these few exceptions, however, the overall results of the typologically unrelated items (see Table 9) reflect the degree of difficulty to infer meaning of unknown words when they are entirely unrelated to the learners' background languages.

Table 9. *Descriptive Statistics (Translation Task: Typologically unrelated words)*

	Min	Max	Mean	SD
Correct translation (8)	0	5	.82	1.29
• *Transfer (other knowledge) (8)*	0	2	.24	.56
• *Prior knowledge (8)*	0	4	.58	1.07
Incorrect translation (8)	0	8	5.00	2.16
No translation / omission (8)	0	8	2.18	1.96

In order to provide a more detailed account of the outcome of the translation task, we also listed the numbers of correct and incorrect translations, as well as their absence for each of the test items separately (see Table 10). Whereas the German-English cognates will be discussed in more detail in relation to the response variable, i.e., positive transfer from

English, in the following subsections, there will be no further mention of the other two groups of words, given that they are not of immediate interest for the present inquiry. However, our preliminary analyses suggest that subsequent investigations into the particular kinds of reasoning exhibited in the verbal protocols of the Latinate borrowings and the typologically unrelated words may be a fruitful endeavour. Learners' reflections on similarities between entirely unrelated words such as *Messer* and *mesa* (Spanish for *table*) or *Kleid* and *kite* (see pilot test) might provide further insights into crosslinguistic interaction in multilinguals.

Table 10. *Descriptive results of the translation task per test item*

	German test item	Translation equivalent	Correct translation		Incorrect translation	No translation
			Positive transfer	Prior knowledge		
German-English cognates	Apfel	apple	47	13	3	3
	Feld	field	31	0	21	14
	Fisch	fish	60	5	0	1
	Freund	friend	39	17	9	1
	Sohn	son	36	20	8	2
	Schiff	ship	11	0	44	11
	Blut	blood	6	2	48	10
	Herz	heart	9	1	38	18
	Wasser	water	29	7	19	11
	Kirche	church	17	3	35	11
	Haar	hair	10	0	24	32
	Licht	light	17	0	24	25
	Brust	breast	1	1	40	24
	Wetter	weather	25	2	25	14
	Maus	mouse	59	2	4	1
	Grund	ground	46	0	15	5
	Milch	milk	52	8	2	4
	Buch	book	25	15	16	10
	Sturm	storm	55	2	5	4
	Kuh	cow	3	2	46	15
	Socke	sock	55	0	8	3
	Knie	knee	56	0	6	4
	Nachbar	neighbor	23	0	28	15
	Gast	guest	13	0	38	15
	Sonne	sun	23	15	23	5
	Stein	stone	3	1	32	30
	Bäckerei	bakery	30	4	23	9
	Regen	rain	2	0	40	24
Latinate borrowings	Meer	mer (fr.) / mar (sp.)	39	1	18	8
	Ferien	férié (fr.) / feria (sp.)	23	1	23	19
	Möbel	meuble (fr.) /muebla (sp.)	34	2	25	5
	Fenster	fenêtre (fr.)	15	2	39	10
	Krawatte	cravate (fr.)	61	0	3	2
	Tasse	tasse (fr.) / taza (sp.)	61	3	1	1
Typologically unrelated	Berg	mountain	4	4	48	10
	Messer	knife	0	8	46	12
	Lied	song	4	10	43	9
	Zeitung	newspaper	2	7	38	19
	Kleid	dress	0	2	38	26
	Luft	air	6	0	45	15
	Pferd	horse	0	6	30	30
	Spiegel	mirror	0	1	42	23

5.1.2 Positive lexical transfer from English as a translation strategy

On the basis of this word translation task from German (L3) into French (L1), positive lexical transfer was operationalized as a correct translation that was explicitly or implicitly related to an English cognate. Given that only 28 of the 42 test items were Germanic-based words that had an English cognate, the maximal possible score to be obtained on this task was 28. First, the descriptive results will be presented and briefly discussed with respect to our first research question that focused on whether or not transfer from English was actually used as a strategy among the targeted population to infer meaning of unknown words in German. Then the focus will be on the test items themselves and the frequency with which they have been transferred, both in relation to similarity ratings between the English and German translation equivalents as well as word frequency in each language.

The first analysis of the think-aloud protocols (TAPs) showed that all of the participants in our sample resorted to lexical transfer from English on a number of occasions to translate unknown words in German into their native language French. More precisely, the overall test results revealed a mean transfer rate of 11.86 with a minimum of 4 and a maximum of 21 out of 28. With reference to our first research question, a student t-test (see Tables 11 and 12) was used to ascertain whether learners of German (L3) significantly resorted to lexical transfer from English (L2) in the translation task.

Table 11. *One-Sample Statistics (Transfer from English)*

	N	Mean	SD	Std. Error Mean
Transfer from English	66	11.86	3.47	.43

Table 12. *One-Sample Test (Transfer from English)*

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Transfer from English	27.79	65	.000	11.86	11.01	12.72

With a mean score of 11.86, a standard deviation of 3.47 and a confidence level of 95%, the confidence interval was +/- 0.86. That is to say that the mean score for the true population would range from 11.01 to 12.72, which is significantly different from zero (see Table 12, *p* < .001). In other words, our initial observation that positive lexical transfer was highly probable to occur in the targeted population could thus be confirmed. However, neither did all of the participants use this strategy to the same extent, nor did they do so for the same test items. What follows is a more detailed account of the distribution of transfer rates among the test items.

5.1.3 Overview of transferred items

Not surprisingly, some items were translated on the basis of transfer from English more often than others. Figure 2 provides an overview of the 28 test items in decreasing order of frequency with which they have been transferred from English throughout the sample. More precisely, the bar diagram below indicates how many out of the 66 participants in our sample translated each of the listed test items by resorting to positive lexical transfer from English.

Number of participants: 66
Most frequently transferred: *Fisch* (N = 60)
Least frequently transferred: *Brust* (N = 1)
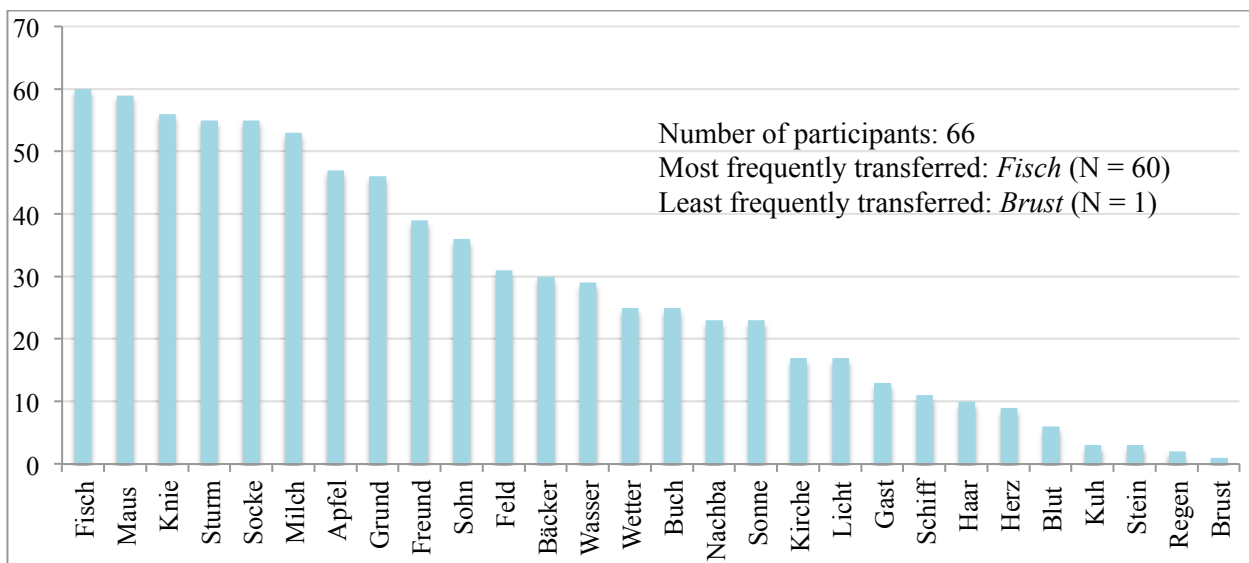
Figure 2. *Transfer rates per test item*

Before comparing transfer rates between participants and trying to determine possible factors that may lead L3 learners of German to resort to positive transfer from English, it was important to consider the relative bias induced by the test items themselves. Namely, if several items were transferred by the majority of the participants while others could only rarely be related to an English cognate, this tendency was, at least partly, related to the degree of similarity between the German target item and its English translation equivalent. For instance, the two most frequently transferred items were *Fisch* and *Maus*, which are homophones with their English translation equivalents *fish* and *mouse*, respectively. It goes without saying that the correspondence between these two pairs of cognates is more easily noticeable than, for instance, the word *Licht* [lıçt] in relation to its English counterpart *light* ['laıt]. Given that the test items had been extracted from a study in which the German and English translation equivalents had been rated for similarity, we tried to relate the order of frequency with which our German items had been transferred from English to the degree of similarity of cognate pairs, as established in Friel and Kennison's (2001) study.

5.1.3.1 Transfer and similarity ratings

As briefly explained in section 4.3.2, the participants in Friel and Kennison's (2001) study who were asked to rate the similarity of the German-English cognate pairs were monolingual speakers of English who did not know any German. Translation pairs were

rated on a scale from one (1) to seven (7), indicating a continuum from very low and to very high similarity (see Friel & Kennison, 2001, pp. 255-256). The task was conducted in two conditions, either with or without a pronunciation guide. Figure 3 provides an overview of the similarity ratings of the 28 translation pairs evoked by the German words used in our translation task. As stated previously, we only chose items that displayed a relatively high degree of similarity to their English cognate, i.e., at least 4 out of 7 in either of the two conditions, with the exception of *Regen* and *Brust*. Namely, as discussed in the previous chapter, the rating for *Regen* in relation to *rain* was only around 3, while that of *Brust* and *breast* was not available (see section 4.1.2). For each item, the bar on the left hand represents the mean rating by the group who had access to a pronunciation guide, while ratings based solely on the orthographic representation of the cognate pairs are represented on the right. Again, the items are listed in decreasing order of frequency with which they have been transferred in our study. In comparison to the previous figure, it is thus possible to observe whether the degree of similarity between the German-English cognate pairs may be an indicator of transferability. Given that our participants had access to both written and auditory representations of the test items, the linear trend shown in Figure 3 traces the tendency of similarity ratings of those participants in Friel and Kennison's study who could also rely on both spelling and sound.
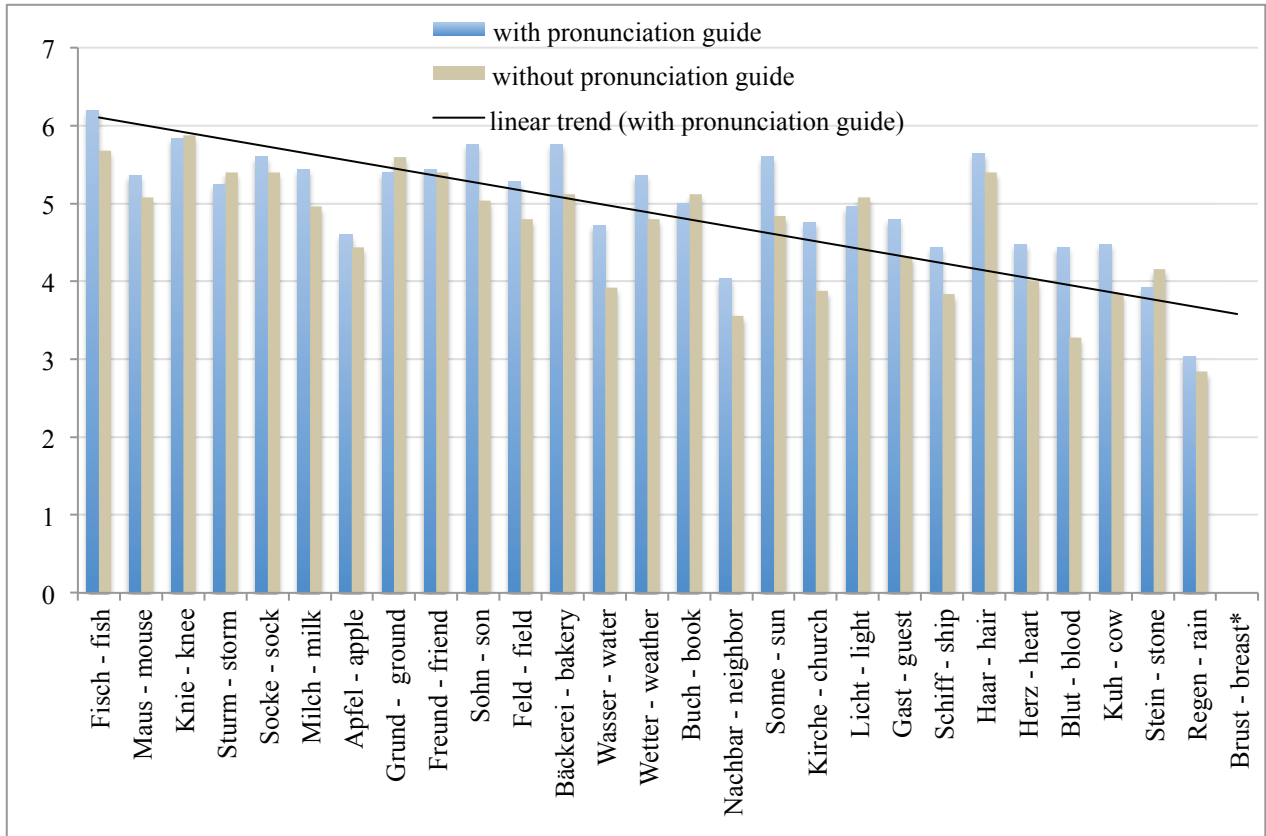
Figure 3. *Similarity ratings (see Friel & Kennison, 2001)*

As illustrated in Figure 3, the similarity ratings established by Friel and Kennison (2001) indicate a tendency comparable to the order with which the German items were transferred from English in our study. That is, the more the given target item is perceived as being similar to its English cognate, the more it is transferred on the basis of this similarity relation. More precisely, statistical analyses (see Table 13 below) revealed that the relationship between similarity ratings and transfer rates was strongly correlated ($r = .67$) and highly significant ($p < .001$).

Table 13. *Correlations (Similarity ratings / Transfer rates per item)*

		Similarity rating with pronunciation guide	Transfer rate per item
Similarity rating with pronunciation guide	Pearson Correlation	1	.67[**]
	Sig. (2-tailed)		.000
	N	27	27
Transfer rate per item	Pearson Correlation	.67[**]	1
	Sig. (2-tailed)	.000	
	N	27	27

Note. ** $p > 0.01$.

However, it seems important to note that a number of translation pairs that were rated close to 6 on the similarity scale have not incited a large number of participants of our study to resort to transfer from English to translate the given word in German. For instance, *Sonne* was translated correctly based on *sun* by 23 of our participants and *Haar* only by 10 of them. This indicates that the degree of similarity does not necessarily determine the frequency of transfer for a given cognate pair. In other words, there is reason to believe that inter-individual variability in the transfer rates is not merely attributable to the task itself.

5.1.3.2 Transfer and word frequency

Since information on word frequency was also available for all of our test items and their English cognates, except *breast*, we listed them in the same order as in the previous graphics and included frequency counts per million for German and English words, respectively.
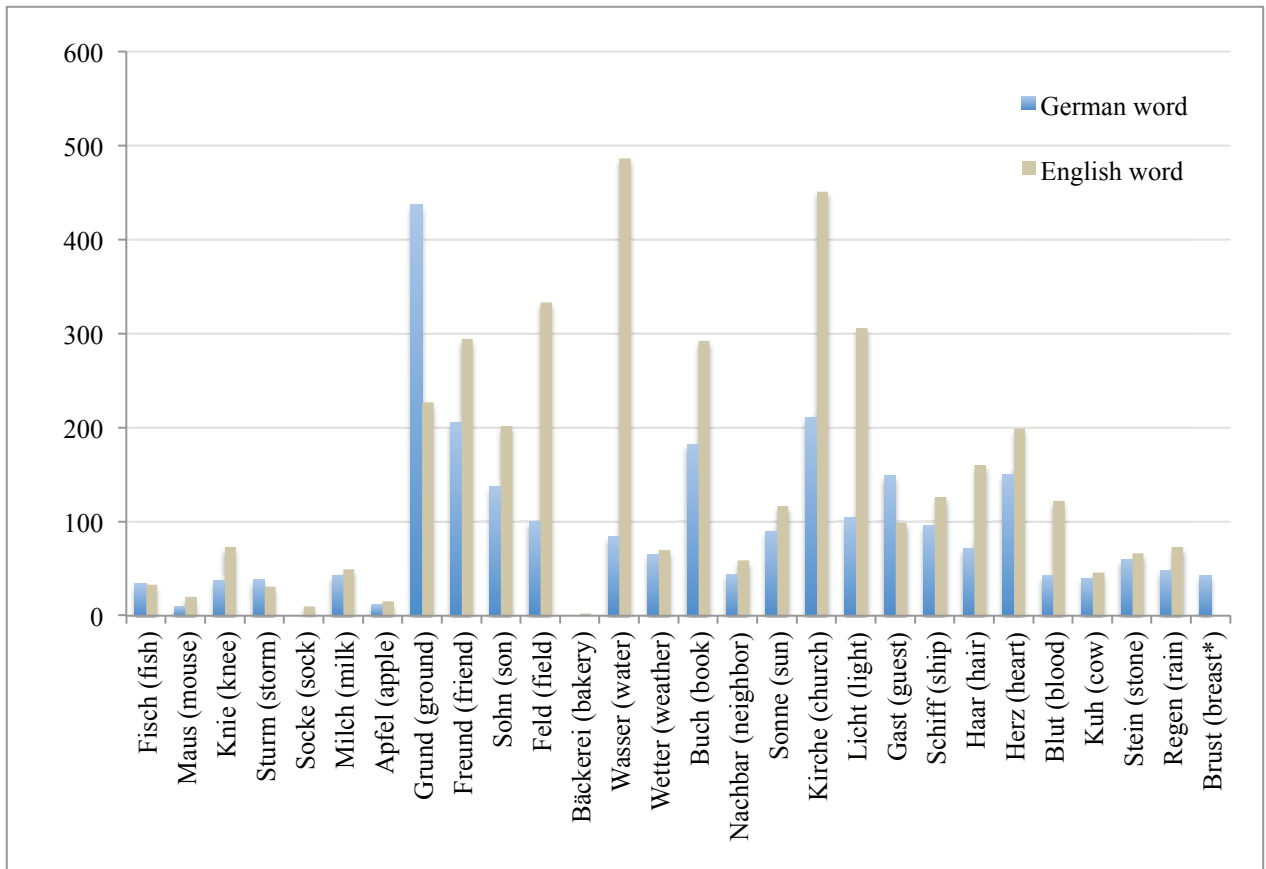
Figure 4. *Word frequency per million (see Friel & Kennison, 2001)*

As evidenced in Figure 4, word frequency in either of the languages does not seem to be related to transfer rates. For instance, the six most frequently transferred items in our study are not particularly frequent, neither in German nor in English. Precisely, Friel and Kennison listed fewer than 50 instances per million for *Fisch*, *Maus*, *Knie*, *Sturm*, *Socke* and *Milch*, as well as their English translation equivalents (except *knee* = 73). In turn, high frequency words in English such as, e.g., *friend* (294), *field* (333), *water* (486), *church* (451) or *light* (306), were not used as a basis of transfer very often. Arguably, the difficulty of identifying *Kirche* with *church* or *Licht* with *light* may be due, at least partly, to the fact that both these German words have a specific sound, namely [ç], which does not exist in the other languages known to our participants. However, *Wasser* [ˈvasɐ] and *Feld* [felt], which do not have particularly complex phonetic patterns, were only related to their English cognates by less than half of the participants.

In sum, it appears as though word frequency did not affect the extent to which individual items were transferred from one language to the other. In fact, there could have been a tendency for high frequency words from English to be activated more easily in the translation process from German to French. Moreover, for the German test items, this characteristic may have increased the chances that learners had already seen or heard the word and thus could have triggered meaningful associations with related words more easily. However, no relationship between word frequency in either language and transfer rates could be established in our sample.

This first subsection was devoted to the results of our translation task, which revealed that all the participants resorted to positive transfer from English, but to different extents. Moreover, the frequency with which each individual test item was transferred from English throughout the sample was analyzed in relation to the degree of similarity between the German-English translation pairs and to the frequency of the respective words in both languages. Whereas no relationship was found for word frequency, similarity ratings were significantly correlated with the transfer rates per test item. Even if a number of cognate pairs rated as highly similar were not recognized as such by our participants or simply not used as basis for translation, it must still be acknowledged that the task itself induced a certain bias, given that some items were in fact more easily transferrable than others. However, the variability among participants, regarding both the number of transferred items and the kind of reasoning used to translate the mostly unknown words from German into French, suggests that individual differences pertaining to cognitive and language-specific abilities might determine this type of crosslinguistic processing. The focus of the next section will be on the manifestations of MLA, such as evidenced in the results of the THAM-3 as well as in the levels of metalinguistic analysis displayed in the TAPs.

5.2   Metalinguistic awareness in L3 learners

When elaborating on the conceptual framework of the present study, a person who is metalinguistically aware has been defined as having *the ability to deliberately direct his or her attention to the structural features of language, and thus to manipulate its components* (see section 2.3.1). The primary instrument to collect data on MLA was a shortened version

of the THAM-3. However, as a second step toward a finer-grained analysis of levels and types of metalinguistic processing, the participants' verbalizations recorded during the translation task were also rated for levels of awareness with respect to crosslinguistic comprehension strategies. Quantitative and qualitative analyses of both types of data will be presented and discussed below.

5.2.1 Levels of awareness exhibited in the THAM-3

The answers to the different questions of the THAM were analyzed by two independent raters based on the coding procedures suggested by the authors of the test (Pinto & El Euch, 2015). Among the raters were the main researcher, an applied linguist and native speaker of German with near-native proficiency in both English and French, and a French-speaking adult enrolled in a TESL program at Laval University. Along with a solid background in linguistics, the second rater was highly proficient in English and had some knowledge of Spanish. Levels of MLA were established on the basis of two partial test scores, representing a linguistic (L) and a metalinguistic (ML) dimension, respectively, which were then combined to a total score. An overview of the levels attained for each type of analysis is provided in the table below. Descriptive statistics are listed for partial and total scores (see Table 14).

Table 14. *Descriptive Statistics (THAM: L, ML and total scores)*

	Min	Max	Mean	SD
THAM_Lscore (16)	4	16	10.79	3.17
THAM_MLscore (32)	0	22	7.58	4.76
THAM_Total (48)	6	37	18.36	7.09

Whereas the maximal score of 16 for the L dimension of the test was attained by some of the participants, none of them came close to the maximal score for the ML dimension, which was of 32 points. More precisely, the highest ML score was of 22 and the mean only of 7.58 thus representing less than 25% of success on the justification part of the test. Prior to a more detailed description of levels and types of analysis across the different sections of the test, we will briefly document and discuss the distribution of L and ML scores in relation to each other.

138

As previously stated, the L dimension of the test was taken to represent an initial stage of awareness, where learners show the ability to distinguish particular meanings or patterns of language. For each test item, the successful identification of a specific relation between linguistic features, as in the *comprehension* and *figurative language* parts, or that of a grammatical error, as in the *acceptability* part, was the first step towards metalinguistic analysis. In other words, the relative depth of metalinguistic analysis depended on whether the given characteristic was correctly identified in the first place. Hence, a rating of L = 1 was a prerequisite for a rating of ML ≥ 1. This is clearly reflected in the scatterplot of the results according to these two dimensions of the test (see Figure 5). Namely, it was possible to reach a high L score but remain on a relatively low ML level, as illustrated by the presence of dots in the top left corner, whereas high ML scores were only attained in combination with high L scores.
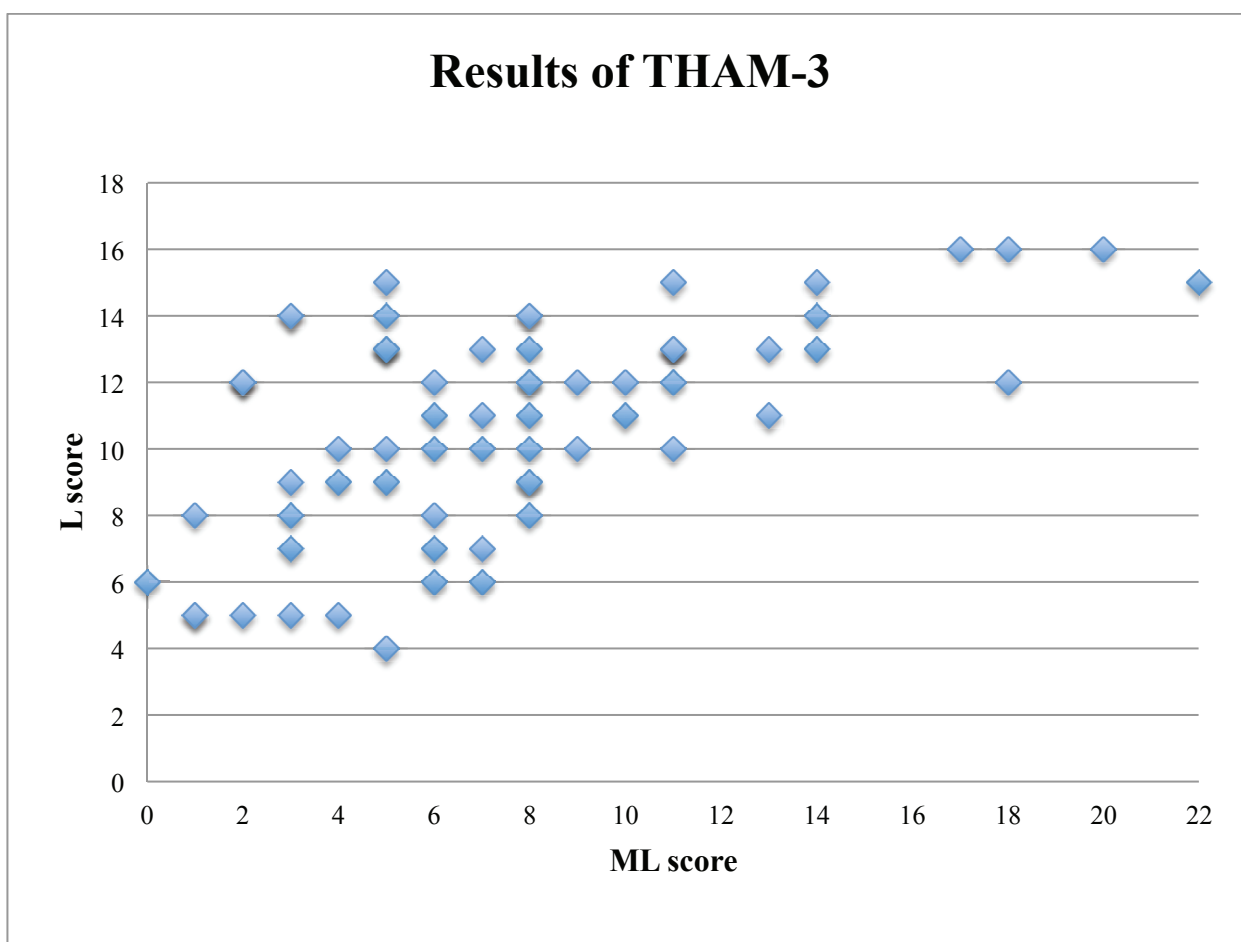
Figure 5. *Results of the THAM (Distribution of L and ML scores)*

However, the distribution illustrated in Figure 5 also suggests a positive correlation between the two scores, which was confirmed by a Pearson correlation ($r = .58$, $p < .01$). This relationship is coherent with our conception of MLA, in which the ML dimension of the test is taken as the extension of the L dimension, while both types of analysis are considered to be a part of the construct. Furthermore, the distribution of levels of analysis for the two dimensions in relation to one another was not consistent across test items, especially with respect to the different types of questions.

5.2.1.1 Levels and types of analysis across test items

Throughout the sample, no notable tendencies could be observed for the type of analysis rated as "linguistic". In other words, there were no test items that caused particular trouble to a large number of participants, as regards the recognition or "noticing" of the targeted features. Still, the number of answers rated $L = 0$ were slightly more frequent in the *acceptability* part of the test. That is, an average of 25 participants (roughly 38%) failed to detect the errors in the short text, whereas the different types of relations between word meanings and functions presented in the other questions were correctly identified by a larger majority, and failed by less than 10 participants, on average (roughly 15%). Having had a chance to observe the participants as they were working on the test, there is reason to assume that this was not necessarily due to the difficulty of detecting errors in the text, but rather to a lack of interest or motivation for this part of the test. A closer look at the paper copies revealed that many participants had not put much effort into it. Precisely, errors were often only underlined and/or annotated in the text, but not necessarily corrected. However, the correction was necessary for an answer to be rated $L = 1$, which partly explains the rather high rate of fails on the linguistic dimension of the error correction part. All in all, the performance on the L score did not seem to depend on the difficulty of specific test items.

However, the "metalinguistic" type of analysis, which required a reflection about the observed pattern and a written justification of its functioning, was clearly more frequent with the items of the *comprehension* (#1-3) and the *figurative language* (#4) sections, respectively.

140
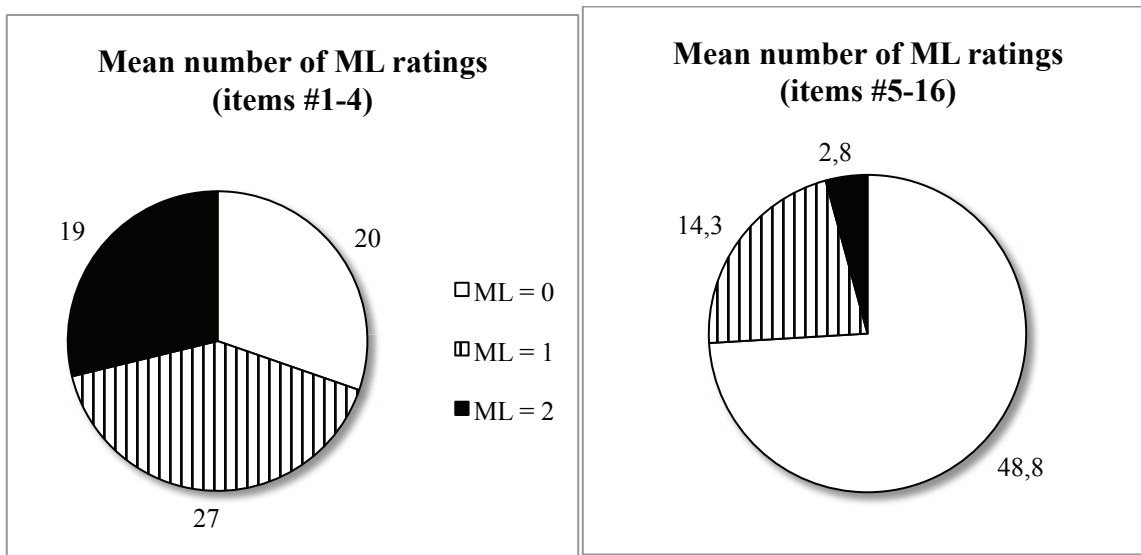
Figure 6. *Distribution of ML ratings across test items (THAM)*

As illustrated in Figure 6, slightly over 40% of the participants (mean = 27) reached at least an intermediate level (ML = 1) of metalinguistic analysis when answering the first four questions of the test. Nearly a third of them (mean = 19) even analyzed these items thoroughly (ML = 2), thus demonstrating that they understood the underlying semantic or pragmatic functions of the given expressions, as opposed to only 4% (mean = 2.8) who reached the maximal ML score in the *acceptability* part of the test. In turn, those who showed a lack of metalinguistic analysis (ML = 0) for these twelve grammar-based items represented 74% of the group (mean = 48.8). This differential distribution of levels of analysis between the different parts of the test can, at least partly, be explained by the nature of the questions themselves. A number of examples from each part of the test will be used to illustrate this point.

Below, examples of the distinct levels of metalinguistic analysis are presented in ascending order (ML = 0 / ML = 1 / ML = 2) for the first and the fourth test item, respectively.

Item #1: *Les phrases suivantes comportent des relations qualitatives. Indiquez s'il s'agit du même type de qualité dans les deux phrases et expliquez ce qui vous permet de répondre ainsi. [English: The sentences that are going to be presented to you deal with qualitative relations. You will need to evaluate whether both sentences deal with the same type of quality or not, and to give reasons for your answer.]*

141

- **Il a une volonté de fer.**
- **Il est borné.**

(1) Les deux phrases sont complètement différentes. Dans la première phrase, « une volonté de fer » la personne s'arrangera pour arriver à ses fins, mais qu'elle s'instaurera des limites malgré tout. Cependant, dans la deuxième phrase, « il est borné », la personne arrivera à ses fins peu importe ce qui en coute (sic), donc elle ne s'instaurera aucune limite. (ID 10: L = 1, ML = 0)

(2) Ce n'est pas le même type de qualité. Une volonté de fer veut dire qu'il est prêt à tout pour y arriver et qu'il ne se découragera point. Tandis que borné renvoit (sic) à quelqu'un qui est restreint mentalement. Une personne qui ne veut pas voir clair ou comprendre. (ID 35: L = 1, ML = 1)

(3) Ce n'est pas le même type de qualité. La première suggère une qualité positive. Quelqu'un avec une volonté de fer ne se laisse pas influencer par les autres et suit ses propres valeurs. Quelqu'un de borné, par contre, refuse de voir la raison lorsqu'il est confronté à un argument qui se tient. Peu importe ce que les autres diront, il aura toujours raison. Celle-ci est une qualité négative. (ID 12: L = 1, ML = 2)

Item #4: *Vous trouverez ci-dessous un slogan publicitaire tiré d'une annonce publicitaire. L'objet auquel il fait référence est indiqué à côté du slogan, entre parenthèses. Indiquez comment ce slogan devrait être compris et justifiez l'utilisation des mots qui le composent. [English: You will be given an advertising slogan taken from an advertisement. The object to which it refers is indicated next to the item. You will need to say how it should be interpreted and to justify the meaning of the words.]*

- **N'eau fatigue, n'eau stress. (eau minérale)**

(4) Nous pouvons comprendre qu'avec ce produit, nous pouvons éviter la fatigue qu'amène la déshydratation et que nous pouvons prendre le temps de relaxer, de nous calmer en prenant un peu d'eau et donc, d'éviter le stress. (ID 46: L = 1, ML = 0)

(5) C'est un jeu de mot sur le mot no (anglais) et le fait qu'on vend de l'eau. On doit comprendre no, fatigue, no stress, pas de fatigue ni stress. L'eau minérale qu'ils vendent est sensée vous rafraîchir, vous relaxer. (ID 32: L = 1, ML = 1)

(6) Le slogan est à la base une phrase en anglais mais est fait pour compris (sic) par des francophones. « N'eau » est homophone de « no » ; « fatigue » et « stress » se retrouvent aussi bien en français qu'en anglais. Le but est de faire un jeu de mots mettant en valeur le mot « eau » et de passer aux consommateurs le message que grâce à leur eau minérale, ils vont pouvoir reprendre de la vigueur (no fatigue) et relaxer (no stress). (ID 5: L = 1, ML = 2)

As evidenced by these examples, the type of metalinguistic analysis heeded in this part of the test was mainly based on semantic aspects of the language, the description of which did not necessitate the use of any specific grammatical vocabulary. Rather, depth of analysis depended on the degree to which the answer covered all or only part of the aspects the expression consisted of. For instance, the qualitative difference between the answers in (2) and (3) was that the former only paraphrased the two expressions, whereas the latter also specified the opposition of a positive vs. negative a quality. As for the advertising slogan, the answers given as examples under (5) and (6) both mentioned the wordplay on "no fatigue, no stress", yet only the latter actually explained how the elements of the slogan are arranged to create the pun. The reason for a rating as ML = 0 was a lack of sufficient explanation. That is, in the example (1), the difference between the two qualities was broken down to "s'instaurer des limites" or not, whereas this aspect is not the main semantic trait that differentiates "volonté de fer" from "borné". As for the advertising slogan, the participant who gave the answer in example (4) did not elaborate on the use of any particular word or sound to create the effect. However, as illustrated by the other examples, there was no need for explicit knowledge of grammatical categories to demonstrate metalinguistic awareness in these comprehension-based test items.

As previously stated, the distribution of the three levels of analysis was quite different in the *acceptability* section (#5-16), where the detected errors had to be categorized (the appropriate grammatical category named) and contextualized (the context of use explained) in order to reach the maximal ML score (ML = 2). When only one of the two analytic steps

143

was taken, the level was rated as intermediate (ML = 1). As shown in the pie chart above, the vast majority of the participants did not justify their answers sufficiently or did not justify them at all (ML = 0). As briefly mentioned with respect to the L score for this section, the tendency to ignore the justification part altogether may be related to a lack of interest or motivation for this kind of exercise. However, the instances where an explanation was attempted often implied tautologies where the error in question was simply qualified as inappropriate, yet no justification was given. Even when an item was partially analyzed, either by categorization or by contextualization, this kind of metasyntactic or grammatical analysis proved to be more difficult across the sample, as evidenced by the relatively small number of participants who reached an intermediate ML level (mean = 14.3). Finally, those who reached the maximal ML score for any of twelve grammatical errors to be analyzed in the *acceptability* part were the exception. To provide a more comprehensive account of the type of metalinguistic analysis heeded in this part of the test, examples of each of the three levels are provided for two of the twelve items. Given that they were part of a text, we listed the respective sentence with each of the target items underlined. Errors had to be identified and corrected, and each correction justified.

Item #7: Il a regardé par la fenêtre vers la direction nord le petit triangle de désert <u>de lequel</u> les falaises ne cachaient pas.[50]

> (7) de lequel → que: « de lequel » est une traduction de l'anglais « which » ce qui ne se dit pas en français (ID 15: L = 1, ML = 0)

> (8) ajout du déterminant « que » au lieu de « de lequel » car il désigne le petit triangle de désert (ID 37: L =1, ML = 1)

> (9) de lequel → que: entre triangle du désert et falaises, il faut un coordonnant pour faire la liaison, par un pronom (ID 12: L = 1, ML = 2)

Item #13: Ils entraient sous différents prétextes […] rapportant des nouvelles insignifiantes: […] qu'environ dix soldats en congé <u>devront</u> déjà être de retour […]

> (10)« dix soldats en congé <u>devrait</u> déjà » ils ne sont pas encore là, le verbe ne peut pas être au présent (ID 65: L = 0, ML = 0)

---

[50] Since the errors in the text are specific to the French language, an English translation would not have rendered the meaning and scope of the error. The same applies for the following example taken from this section, which is why English translations seemed inappropriate.

(11) devraient [possibilité du passé = conditionnel] (ID 58: L = 1, ML = 1)

(12) devront → devraient: conditionnel, il s'agit d'une supposition (ID 27: L = 1, ML = 2)

The examples of answers for both these test items point to the fact that grammatical or syntactic analyses that might seem straightforward to a linguist proved to be rather demanding for our participants. The type of justification that was most frequent for the totality of the twelve items of the *acceptability* section, was the one exemplified in (7) and (10), respectively. To be more specific, whether or not the error had been correctly identified and corrected, most of the explanations were too vague to qualify as metalinguistic. For instance, to state that a given structure does not exist in French, such as illustrated in (7), or that the use of the present tense is inappropriate, as in (10), does not justify the correction. Rather, it is a circular argument that remains self-referential. The second most common type of answer (mean of 14.3, roughly 22% per item) was categorized as ML = 1, such as illustrated in (8) and in (11). The former example shows an instance where the context of use of the appropriate structure (contextualization) was explained but not categorized, or wrongly categorized. Namely, "que" is not a determiner but a pronoun. In the latter example, the categorization is appropriate ("devraient" is a conditional), while the contextualization is not. In this case the conditional does not express a 'possibility of the past' but rather a 'supposition', as exemplified in (12). Throughout this section, the maximal number of participants who correctly justified their answers using both contextualization and categorization was seven for a single item whereas for the other eleven items, scores of ML = 2 were attained by only one to four out of 66 participants (mean = 2.8, see Figure 6 above).

The fact that a large number of participants failed to analyze grammatical errors or resorted to tautological statements as justifications could be an indicator of their lack of explicit knowledge of the respective underlying structures. On the other hand, it is possible that they believed this type of pseudo justification was sufficient, given that they had already detected and corrected the error. Whichever of the two interpretations is the more appropriate, our results suggest that explicit grammatical knowledge, as is typically targeted in grammaticality judgment tasks used as a common measure of MLA, represents a

type of metalinguistic analysis that the participants of our study were either not familiar with or otherwise not inclined to engage in.[51] On the other hand, the majority performed rather well on the metasemantic and pragmatic types of analysis, as reflected in their answers to the first four questions. These observations suggest that a more complete picture of MLA is needed in order to investigate the scope of what it means to be "able to deliberately direct one's attention to the structural features of language, and thus to manipulate its components" (our working definition of MLA, see Chapter 2). According to our conception of MLA, these structural features cannot be restricted to correct identifications of grammatical categories and their application. That is to say, the deliberate manipulation of the language one is using can be reflected in different types of analysis in all areas of the language, including semantics, pragmatics, phonetics, etc. Moreover, in line with our previously stated conception of awareness, which is mostly inspired by Schmidt's notion of "noticing" applied to SLA research on MLA by Leow (1997, 2000, 2011), for instance, such manipulation may already be reflected at initial stages of awareness, i.e., where particular patterns or features are merely noticed, but not explicitly analyzed on a deeper structural level. As will be explained in more detail with respect to the coding of the TAPs (see section 5.2.2), awareness at the level of "noticing" was also identified in the verbalizations gathered during the translation task. However, given the nature of the translation task, a further subdivision was made between the categories labeled "intuition" and "noticing", both of which represented this intital stage of awareness, i.e., *noticing*, yet exhibiting distinct levels of explicitness.[52]

5.2.1.2 Inter-rater reliability and final scoring

As previously mentioned, two independent raters evaluated all of the 66 participants on their performance on our version of the THAM-3. Inter-rater reliability for the linguistic (L) dimension of the test was of 99.72 percent. To be more specific, we listed a total of three instances throughout the sample (16 x 66 = 1056) where linguistic answers were not rated

---

[51] Learners having been schooled in an era where communicative language teaching may be considered the "current mainstream" (Williams, 1995, p. 12) may not have had many opportunities to focus on form in their second language classrooms. In turn, many of them might lack the practice of reflecting on grammatical forms in a systematic way.

[52] Associations that were explicitly related to an English cognate or to another similar word that could possibly have fit the context were coded as "noticing", associations without explicit references to crosslinguistic correspondences were coded as "intuition" (refer to section 5.2.2.1 for examples).

identically. Two of these instances referred to the same item, in which the deviation was based on the interpretation of whether or not the advertising slogan (item #4) had been understood by the participant, thus rated as $L = 1$ or 0, respectively. The third instance referred to item #7, in which the erroneous use of "de lequel" had to be corrected (see examples 7 to 9 above). The participant had corrected the error by using "lequel" as a relative pronoun[53], which was accepted as grammatical by one of the raters but rejected by the other.[54] Apart from these isolated cases, the linguistic dimension of the test was not subject to much interpretation, as reflected in the high degree of correspondence between the raters. As for the metalinguistic dimension of the test, the evaluation was less objective, given that they were based on a qualitative analysis of the answers. Even if the authors of the test (Pinto & El Euch, 2015) provided ample documentation as to the interpretation of different kinds of answers, including examples for each level of ML analysis, there were some deviations in the scoring of these answers. Per test item, an average of 5 to 6 answers out of 66 were scored as a different level of ML analysis. Throughout the sample, this represented 8.4 percent of deviation, thus resulting in roughly 91.6 percent of correspondence between the two raters for the ML dimension of the test. After integration of the two types of ratings into a combined score, inter-rater reliability was of 95.6 percent. Figure 7 below illustrates the distribution of identical versus deviant ratings for both dimensions of the test.

---

[53] The participant in question (ID 64) had corrected the errors by rewriting the text. The relevant section went as follows: Il a regardé par la fenêtre en direction nord le petit triangle de désert lequel les falaises ne cachaient pas.

[54] The reason why the second rater rejected this correction as unacceptable was that the coding procedures provided by the authors of the test (Pinto & El Euch, 2015) listed "que" as the correct pronoun. After confirmation that "lequel" was in fact possible in this context, the answer was accepted as an appropriate correction ($L = 1$).

Figure 7. *Inter-rater reliability – THAM (Distribution of identical vs. deviant ratings)*

As outlined in Chapter 4, when elaborating on the implementation of the THAM-3 in our study, the two partial test scores (L and ML) were collapsed to form a combined score of MLA. The maximal score attained on this integrated measure of MLA was 37 out of 46 and the minimum was 6. Figure 8 represents the distribution of the total scores in ascending order throughout the sample.
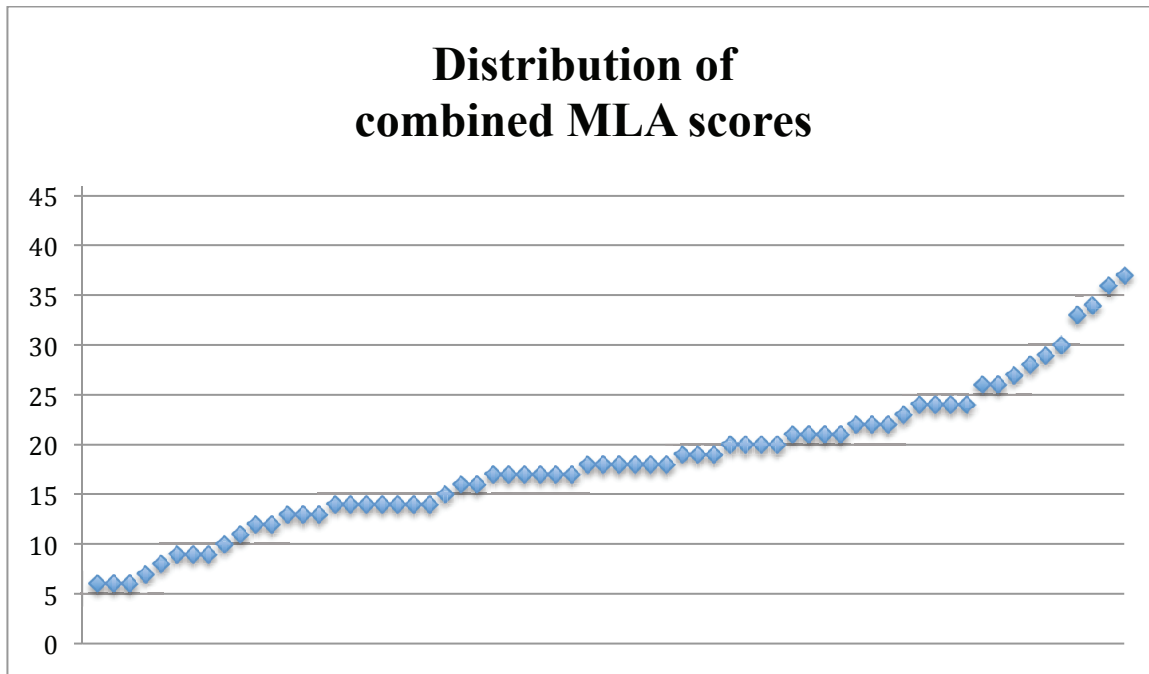
Figure 8. *Distribution of combined MLA scores (THAM)*

On the one hand, this visually supports the evidence that the number of participants who scored above and below the mean, respectively, was nearly identical (mean = 18.36, median = 18). On the other, this gradual distribution confirms the absence of outliers. Having thus presented and discussed the descriptive test results of the THAM-3, the following section provides an outline of the results obtained on the secondary measure of MLA, as exhibited in the TAPs.

5.2.2   Levels of awareness exhibited in the TAPs

As in the evaluation of the THAM-3, the verbalizations produced during the translation task were analyzed by two independent raters – the same who had rated the THAM-3 – who applied the coding procedures described in the last section of Chapter 4, pertaining to data analysis. Levels of MLA were established according to six categories, four of which represented the graded levels of metalinguistic analysis, ranging from *Intuition* (1) to *Noticing* (2), *Meta-Awareness* (3) and *Underlying rule* (4). The other two categories, *Absence* and *Prior knowledge*, were rated zero (0), given that both implied absence of analysis (see discussion on coding procedures above). With a total of 28 test items under observation, the maximal score on this metalinguistic dimension of the task was of 112 (4 x

28). Descriptive statistics of the overall results are listed in Table 15.

Table 15. *Descriptive Statistics (MLA score from the TAPs)*

	Min	Max	Mean	SD
MLA_TAP (112)	17	74	44.41	12.68

Apart from the fact that all participants exhibited at least some sort of analysis, and that none of them attained a total score above 74 (≈ 67 %), these numbers do not provide any insight into the nature of MLA that was exhibited in the TAPs. Hence, the following subsections are devoted to the levels and types of analysis that were evidenced in the verbalizations. The respective ratings will be presented and discussed for transferred and non-transferred test items, respectively.

5.2.2.1 Levels and types of analysis across test items

As previously mentioned, the four levels of metalinguistic analysis that were differentiated in this coding procedure, did not depend on whether or not the respective item had been transferred or not. However, in order to provide a comprehensive account of the relative manifestations of MLA, the overall distribution of levels of analysis will first be presented for transferred and non-transferred items, separately. The two pie charts below (Figures 9 and 10) provide a visual representation of this distribution.

MLA: Transferred items

Legend:
- Non-Transfer
- Intuition
- Noticing
- Meta-Awareness
- Underlying Rule

(58% / 42%) — 506, 242, 26, 9
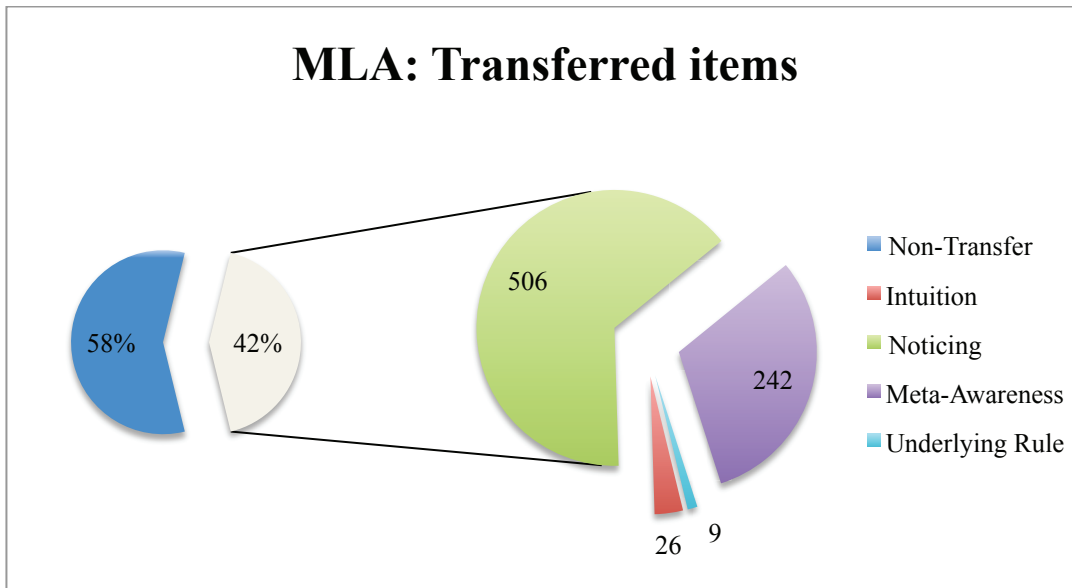
Figure 9. *MLA scores from the TAPs (Number of transferred items per level of analysis)*

Even if our participants significantly resorted to transfer from English (L2) to translate unknown German (L3) words into French (L1), the total number of transferred items throughout the sample represented roughly 42 percent (N = 783). The vast majority of the transferred items were treated at the level of *Noticing* (N = 506). The second most common level was that of *Meta-Awareness* (N = 242), whereas only 9 instances were listed for the highest level of MLA, which was called *Underlying rule*. Finally, there were 26 instances of correct translations that were clearly based on the English cognate, yet with no explicit reference, which was coded as *Intuition*. In proportion to the whole data set, these two categories each represented less than two percent of the 1848 translation attempts (66 x 28 items). As for the non-transferred items, the distribution of levels of analysis across the sample was quite different (see Figure 10).
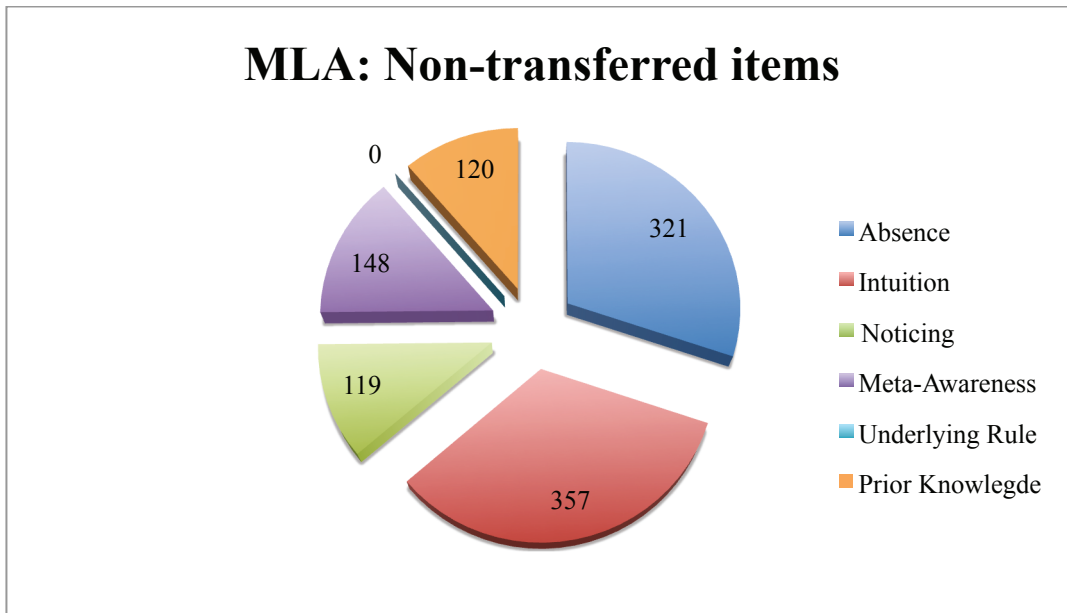
Figure 10. *MLA scores from the TAPs (Number of non-transferred items per level of analysis)*

A closer look at this chart reveals that only 30 percent (321 instances) of the verbalizations that did not reveal transfer from English reflected a total *Absence* of analysis. Rather, the majority of them were qualified as translation attempts that were based on random clang associations, coded at the level of *Intuition* (N = 357). Moreover, and most interestingly, the categories of *Noticing* (N = 119) and *Meta-Awareness* (N = 148) were also applicable when no appropriate connection to an English cognate could be established. In order to illustrate the distinctions between levels of analysis for both transferred (TR) and non-transferred (nonTR) items, a certain number of examples of each level of analysis will be presented for each category.

As previously mentioned, verbalizations coded at the level of *Intuition* were either correct translations that were based on an English cognate but without any explicit reference, as exemplified under (13), or random clang associations with words that could not possibly fit the context, as illustrated under (14).

(13) Euh, *Schiff…* on dirait qu'il y a un mot un peu qui ressemble à l'anglais mais je me rappelle plus trop c'est quoi… on dirait que c'est comme un bateau ou genre un petit voilier ou quelque chose comme ça… parce que *Schiff* c'est comme… (ID 63: TR/I)

(14) *Schiff, Schiff… Schiff,* un peu comme *chiffre* en ang… en français… ou sinon… *Schiff,* euh, en anglais ça me fait penser à rien là, c'est juste c'est *chiffre* en français donc j'irais pour *chiffre…* (ID 25: nonTR/I)

In the first example, the target item *Schiff* [ʃɪf] was correctly translated by *bateau* (French for *boat* or *ship*) based on an association with an English word that was not articulated, but referred to by its resemblance to English. Moreover, the participant used this resemblance as a justification by stating "because *Schiff* is like…" This was one of the rare cases where an item was coded TR at the level of *Intuition* (I) given that the explicit reference to the English cognate was missing.[55] As illustrated in the second example, instances of non-transfer were rated at the same level of analysis when the translation was explicitly related to another word from English or any other language, yet solely based on a clang association that could not logically fit the context. For instance, in the example under (14), the association with *chiffre*, which is often pronounced [ʃɪf] in French, is solely based on correspondence in sound. But most importantly, the translation reveals that the participant failed to attempt an analysis of the sentence that accompanied the item, i.e., *Das Schiff fährt über den Fluss* (English: *The ship travels across the river*). That is, even if the sentence was not supposed to unveil the meaning of the target item, the mere presence of the verb *fahren* (English: *to go/drive/travel*) should have sufficed as falsifying evidence, given that a *shiffre* (English: *number*) cannot possibly go or drive anywhere. Other items with which this type of analysis was particularly frequent were *Blut* [bluːt] associated with *blue*, *Wetter* [ˈvɛtɐ] with *better*, *Sonne* [ˈzɔnə] with *sonner* [ˈsɔneː] (English: *to ring*). In these cases, the

---

[55] It is important to note, however, that there were other cases where a participant would provide a correct translation without elaborating on the source of reasoning. When the answer was given without hesitation, these were precisely the cases in which the researcher would prompt participants to specify whether they had already learned the word (see section 4.4.1 on the data collection procedure of the translation task), in which case the answer would be coded as *Prior Knowledge*. It appears that all of these ambiguous cases, i.e., where a prompt was necessary, were identified, given that all of the remaining instances rated as transfer at the level of *Intuition* (N = 26) contained markers of uncertainty, such as rising intonation, long hesitations, repetitions, sentence-checks, or explicit markers of attenuation such as "probably", "maybe", "I'm not sure", etc. As a consequence, it was assumed that verbalizations of the type "Ça ressemble à lait mais ça doit pas être ça" (ID #47) represented transfer from English at the level of *Intuition*.

lack of logic could simply be related to word class: *blue* and *better* are adjectives, and *sonner* is a verb. In sum, verbalizations that were coded at the level of *Intuition* mostly represented random clang associations that did not result in positive transfer. However, a few instances were listed where this intuitive level of crosslinguistic awareness was related to a correct translation based on the English translation equivalent (see example 13). The two subsequent examples listed under (15) and (16) below should further illustrate the kinds of reasoning rated at the level of intuition, some of which did lead to correct translations (see example 15), while most of them did not (see example 16).

(15) *Sonne… Sonne*, ça on dirait comme, ça sonne pas comme ça là, mais on dirait le mot euh… *Sonne…* on dirait peut-être comme un *soleil*, ça commence pareil, comme un *son*, mais ça sonne pas comme un *son*, un *son*, un *soleil* (ID 20: TR/I)

(16) *Sonne…* euh… *Sonne*, on dirait que ça sonne comme un *son*, du bruit… euh… par l'orthographe… ou *sonner*… quelque chose du genre, quelque chose qui *sonne*, qui fait du bruit… (ID 2: nonTR/I)

For a verbalization to qualify as an analysis at the level of *Noticing*, there had to be an explicit reference to a similar word to justify the translation. It could either be the formal resemblance to the English cognate that was reported and used as a basis for a correct translation, as exemplified under (17) and (19). Or else, the participant may have related the target item to a different word (either English or other) that could equally have fit the context but did not lead to a correct translation, as in (18) and (20).

(17) J'ai … eum… *church*, l'église… *Kirche…* sûrement pas… mais bon… ça me fait penser à ça, église… *Kirche…* ouais, église mettons… (ID 17: TR/N)

(18) Hmm… "dans le *Kirche*"… [chuchote]… hmm… un *cirque* peut-être, c'est ça que ça me fait penser, mais… à cause de *circus* en anglais puis même *cirque* en français… *Kirche…* (ID 29: nonTR/N)

(19) Ça me fait penser à cœur, genre *heart… Herz… Herz… Herz…* hmm… hmm… je sais pas… *Herz…* (ID 26: TR/N)

(20) Herz, ça me fait penser à *hair* en anglais, donc cheveux, mais… je suis pas sûre que ce soit cheveux… (ID 48: nonTR/N)

As illustrated under (18) and (20), words other than the English cognate could be activated

in the process of crosslinguistic consultation and result in incorrect translations. However, the difference with the random clang associations discussed above is that both associations listed under (18) and (20) would have been semantically and syntactically acceptable in the context. In turn, when comparing (17) to (18) and (19) to (20), respectively, depth of analysis was identical for successfully and unsuccessfully transferred items. More precisely, the unknown words were explicitly related to similar items from another language and translated accordingly. The identification of this level of analysis qualified as *Noticing* in our study, was inspired by Leow's (1997) category labeled *Cognitive Change* (CC) evidenced by some kind of behavioral or cognitive change due to the learning experience (see Table 5). Given the relative simplicity of the translation task administered in our study, the learning experience was that of associating unknown vocabulary to familiar words and to make such inferences based on attention that is focused on formal resemblance.

We qualified as Meta-Awareness (MA) all those verbalizations that showed a justification based on internal logic, either referring to specific sounds and/or spellings or to the accompanying sentence. This level was inspired by Leow's (1997) MA category, which implied a report of being aware of the experience, including a justification that was based on a pattern that learners had seen emerge. The first two examples illustrate verbalizations in which phonetic and orthographic features of the target item were explicitly compared to the English cognate and taken as a justification. Instead of simply noticing a resemblance between *Maus* [ˈmaus] and *mouse* (see example 21) and *Sturm* [ʃtʊrm] and *storm* (see example 22), respectively, the similarity relation was justified on the basis of specific formal characteristics.

(21) Euh, ben, en tout cas, à l'oreille ça me fait penser à souris à cause de *mouse,* donc c'est peut-être vraiment souris… à l'écrit, par contre, c'est un peu différent… mais ouais j'imagine que c'est peut-être vraiment ça… (ID 18: TR/MA)

(22) *Sturm… Sturm… sternum* ou *storm* comme euh, tempête. *Sturm, Sturm,* ouais… ça doit être… j'ai encore la tempête dans la tête… *Storm, Sturm,* ça s'écrit pas mal pareil, à part que le « u » deviendrait un « o », puis ouais, j'sais pas… (ID 20: TR/MA)

However, meta-awareness was most often reflected in verbalizations in which the sentence was analyzed and the meaning or function of co-occurring words taken to justify the answer. In the pair of examples listed below, the participants analyzed the sentence, word by word, trying to construct the meaning of the target item based on the context. The difference between the two was that the participant who gave the answer under (23) established the formal correspondence between *Schiff* and *ship*, which confirmed the assumption that the correct answer was *bateau*, as opposed to the answer given under (25) where the participant failed to recognize that resemblance and translated the target item by *pont* (English: *bridge*). However, both justifications were based on the same logic, i.e., the target item had to be something that crossed the river.

(23) *Schiff*… hmm… euh… je dirais que c'est bateau… à cause que… « Das Schiff fährt über den Fluss »... un bateau… traverse le fleuve ou la rivière plutôt… ouais, puis ça ressemble un peu à *ship* mais avec le contexte ça l'aide de savoir que c'est *Schiff*… (ID 59: TR/MA)

(24) Euh… ben, ça ressemble à… je pense que ça veut dire *pont*... parce que dans la phrase ben c'est que *Fluss* euh, fleuve, euh rivière… puis *über* ça veut dire au-dessus fait que c'est quelque chose qui passe au dessus d'une rivière… puis ça pourrait être un pont… puis je vois pas de liens avec d'autres mots… (ID 56: nonTR/MA)

Cross-checks with the sentence were also coded as MA when the analysis led the participant to change their mind about the initial choice, due to semantic or syntactic restrictions they had thus discovered in the sentence.

(25) *Gast*… *Gast*… « Herr Müller ist ein seltener Gast »… *Gast*… ça me fait penser à *gaz* mais c'est probablement pas ça, je sais que c'est pas ça à cause de la phrase parce que « Monsieur Müller est un gaz » ça marche pas… *seltener Gast*… (ID 14: nonTR/MA)

(26) Ça me fait penser à mieux *better*... mais ça a pas rapport parce que c'est un « w » et non un « b », mais il me semble dans un autre démo il y avait *Wetter* puis il me semble ça aurait eu de l'allure avec mieux… mais c'est des noms, mais… euh Monat… nous avons des, des *Wetter*… nous avons… (ID 17: nonTR/MA)

In both these cases, no similarity relation could be established to the English cognate that would have led to the correct translation. Yet even if the initial translation choice would have been coded as a random clang association at the level of *Intuition*, both these verbalizations clearly reflect a deeper level of analysis. Precisely, reasons why the given association was inappropriate in the context were explicitly stated. The participant who gave the answer under (25) relied on semantics, affirming that a person could not be a gas. In the other example (26), the learner reflected on differences in spelling between "w" vs. "b", on possibly related meanings between *Wetter* and *better*, by reference to another learning situation, and finally on word class, by stating that the target item had to be a noun and thus could not be translated by *better*. In sum, to qualify for the level of MA, verbalizations had to reflect the reasoning behind the given association. Based on logical inferences, participants thus demonstrated a conscious manipulation of the language they were using, yet without necessarily being able to name grammatical categories or to formulate underlying rules.

Finally, the highest level of analysis in the present study was inspired by Leow's (1997) *Morphological rule* (MR) category (see Table 5). However, given that the type of reasoning that was needed to translate isolated words did not necessitate any kind of morphosyntactic analysis, this level was labeled *Underlying rule* (UR) and referred mostly to the explicit formulation of regularities in patterns pertaining to spelling and sound. When such a pattern was correctly described and used as a justification, the verbalization was coded UR. Since it was not actually necessary to refer to any of those patterns in order to solve the task, it was not surprising that this level was attained only in some rare cases (N = 9). The examples below illustrate the kind of reasoning that was considered as belonging to this category.

(27) Ah c'est soleil c'est ça soleil… bon ben, *Sonne* c'est soleil c'est euh, je le sais parce que ça ressemble à *sun* mais je suis toujours mêlée avec ça puis *Sohn* le fils mais là je le sais parce qu'il y a un « e » puis c'est ça qui fait la différence là, puis le « o » qui est long là, dans *Sohn* parce qu'il y a le « h »… (ID 58: TR/UR)

(28) Une pomme… « Möchtest du einen Apfel? »… Voudrais-tu une pomme ? C'est comme *apple* en anglais… souvent il y a des mots en anglais que c'est un double « p » ou un « p » puis c'est devenu [pf] en allemand… (ID 5: TR/UR)

157

(29) *Lumière* parce que c'est comme en anglais *light...* souvent dans les mots en allemand là, ça devient « c-h », quand en anglais c'est « g-h »… (ID 5: TR/UR)

In the example under (27), the participant reflected on the similarity between the target item *Sonne* [ˈzɔnə] and its translation equivalent *sun* in English, but also pointed to the possibility of getting confused with another similar word in German, namely *Sohn* [zoːn]. When elaborating on this similarity, the participant further indicated which orthographic markers could be used to differentiate between the distinct [o] sounds in *Sonne* and *Sohn*, respectively. On the basis of this rule of thumb, correspondences between spelling and sound of other German words could be similarly identified. Likewise, the answer given under (28) reflects knowledge of a certain sound shift in the Germanic languages. To be more specific, the participant justified the translation of *Apfel* based on *apple* with reference to a common correspondence between [pf] in German and [p] in English. Whether or not the participants actually knew anything about this shift, it was due to this explicit knowledge of a phonological pattern that the translation could be confirmed. The example under (29) illustrates a similar statement with respect to a correspondence between the spellings *gh* in English and *ch* in German, thus confirming the translation of *Licht* by *lumière*, based on *light*. As reflected by the examples above, this level of analysis only occurred in combination with transferred items. In fact, this is not surprising for the analyses of *Apfel* and *Licht* given that the respective sound correspondence was established in relation to the English cognate. Naturally, the recognition of the given pattern was based on the recognition of the relationship between the cognate pairs. However, the observed pattern exemplified under (27) referred to a sound-spelling correspondence in the German language that was used to distinguish *Sonne* from *Sohn*. The discovery of a systematic orthographical representation (here: postvocalic consonants) of specific sounds in German (here: vowel length) did not depend on the recognition of the respective cognates in English, i.e., *sun* and *son*. In other words, it was possible to distinguish between *Sonne* and *Sohn* based on the orthographical pattern in question, and thus to exhibit awareness of an underlying rule, without recognizing the similarity relation to the respective English cognate. In sum, the highest level of MLA that was identified in the TAPs was also the one that was the least frequent across the sample. On the one hand, this may be taken as an indicator that underlying rules or patterns of the German language itself or of systematic

158

correspondences between German and English were largely unknown to the participants. On the other hand, it must be acknowledged that such depth of analysis was not required to solve the translation task. Hence, it is also possible that participants who were aware of specific patterns did not necessarily apply them, when other translation strategies such as cognate recognition or cross-checks with the sentence, sufficed to make the appropriate inferences.[56] Still, it seemed important to distinguish the level of UR from that of MA by the degree of explicitness with which the structural feature that served as a justification was described.

5.2.2.2 Variability within and across participants

In line with our conception of MLA as a momentary mental activity rather than as an underlying ability (see Chapter 2), our analysis of the TAPs revealed inter and intra-individual variability regarding different types and levels of analysis. As suggested by the examples of verbalizations discussed above, the kinds of answers given for the same test item could vary consistently from one participant to the other. Even if the degree of similarity between German target items and their English cognates was, at least partly, determinant for transfer from English to take place (see section 5.1.2.1), the test items themselves did not appear to affect the kind of reasoning that would be applied as a translation strategy. Three test items were chosen to illustrate this: *Fisch*, which sounds exactly like its English cognate, was transferred by the large majority of the participants, nearly half of them recognized the relation between *Feld* and *field*, while only 10 out 66 did so for *Schiff* and *ship*. From a comparison of the charts referring to *Fisch* and *Feld* below (Figures 11 and 12), it is possible to infer a slight tendency for *Noticing* to occur more often with a high degree of similarity and for *Meta-Awareness* to be associated with a less obvious similarity relation.

---

[56] This issue will be further discussed with respect to the limitations of the present study in the following chapter (see sections 6.3.2 and 6.3.4)
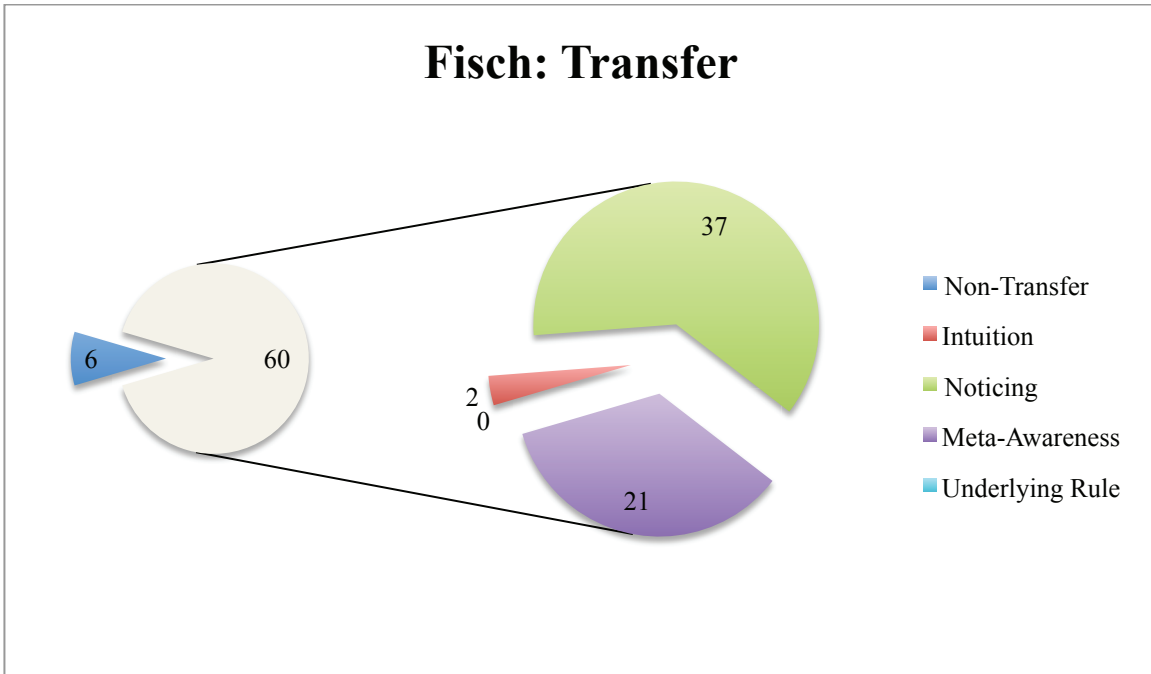
Figure 11. *Number of participants per level of analysis when Fisch was transferred*
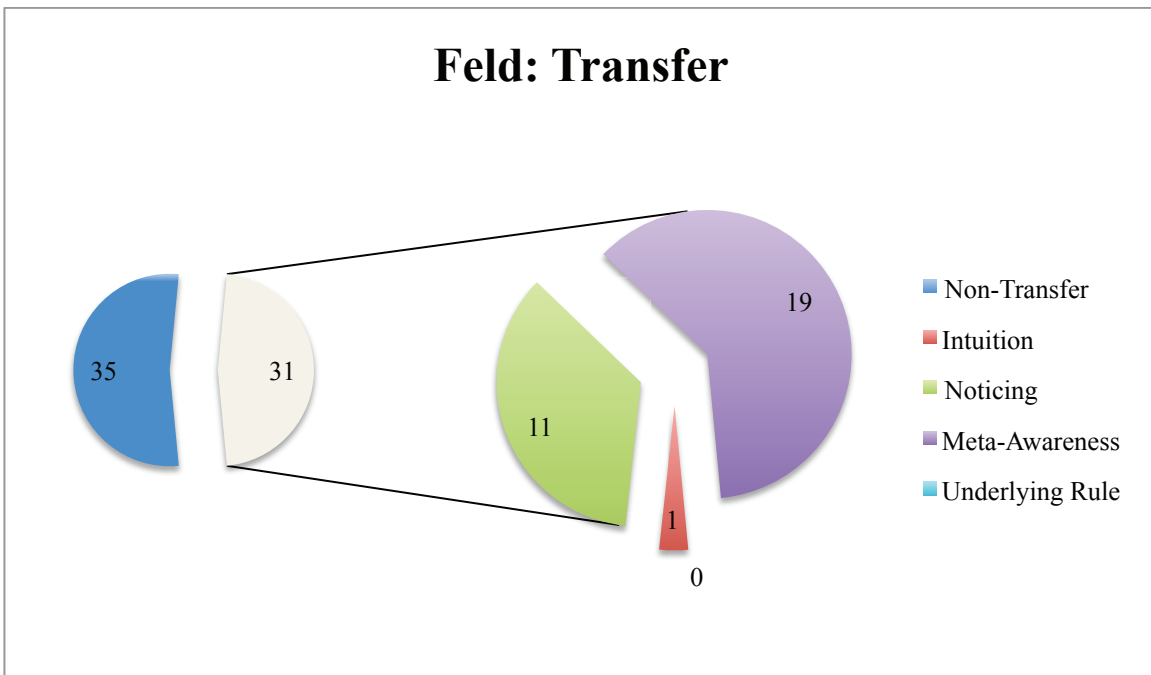
Figure 12. *Number of participants per level of analysis when Feld was transferred*

If such a tendency exists, it could be confirmed by looking at the kind of reasoning exhibited in the verbalizations of those who translated *Schiff* on the basis of *ship*. Namely, when it was transferred, *Schiff* was most often analyzed at the level of MA (see Figure 13).

At the same time, this level of analysis was also attained when *Schiff* was not transferred (see Figure 14).
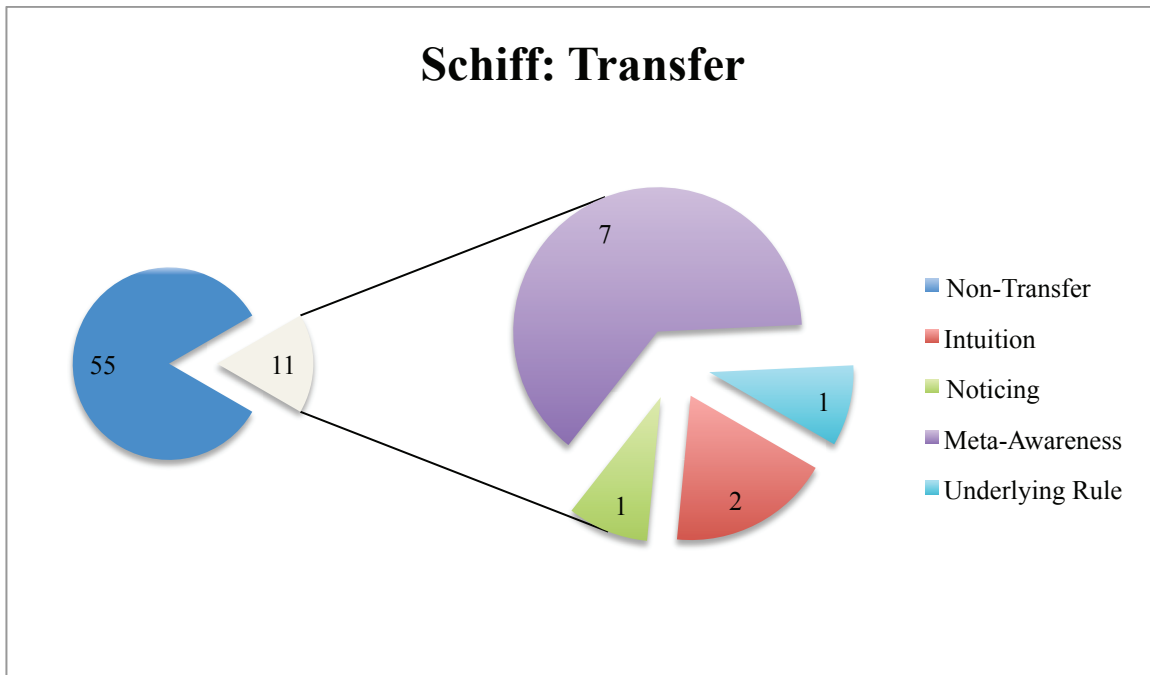
Figure 13. *Number of participants per level of analysis when Schiff was transferred*
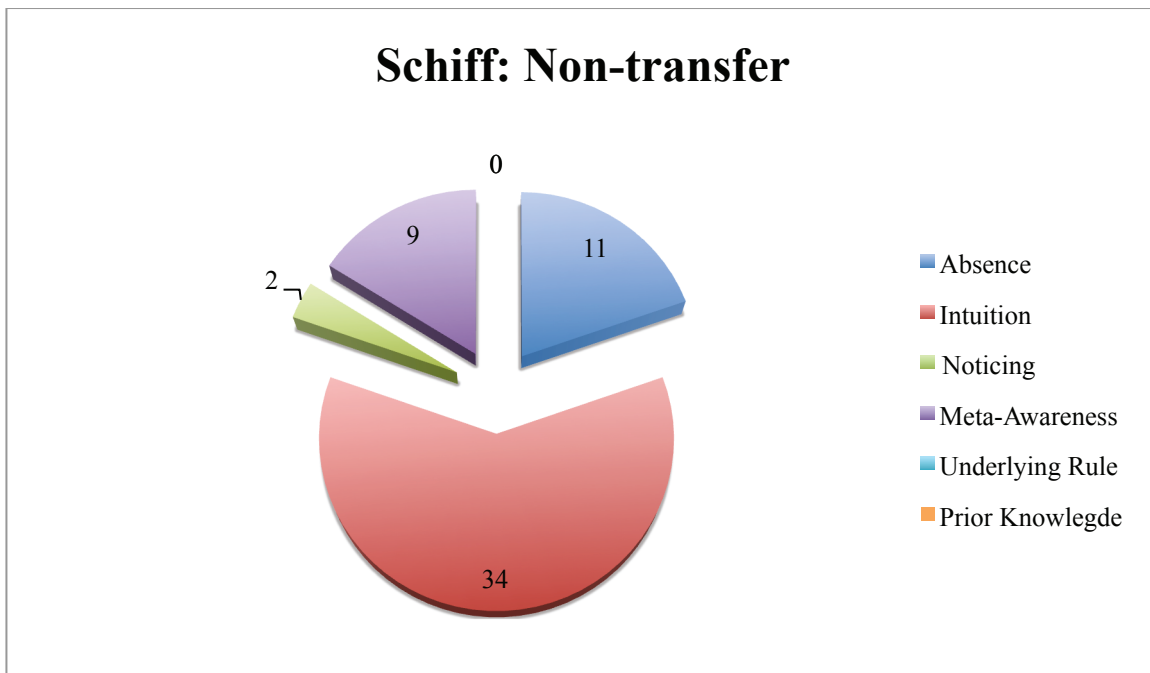
Figure 14. *Number of participants per level of analysis when Schiff was NOT transferred*

These examples of item-wise distributions were used to illustrate the fact that the items themselves did appear to trigger specific types of analysis. In other words, participants interpreted the target words in different ways and by means of different types of strategies. Most interestingly however, this variability of levels and types of analysis could also be observed across the different verbalizations of the same individual. More precisely, when several items were analyzed thoroughly, it might have been expected that other items would be analyzed and manipulated with the same attention and care by the same participant. However, a number of counterexamples indicated that the respective levels of MLA that were manifested in the TAPs could vary consistently across the task. The three verbalizations listed below are from the same participant (ID 53) and should illustrate such intra-individual variability.

(30) Euh, ça ça doit être un mot de la même… ben, pour désigner dans la famille… ou un ami comme *friend* peut-être… comme en anglais… *Freund* parce que dans la phrase, ça veut peut-être dire « Peter, c'est mon ami » comme en anglais, ouais comme j'ai dit, *friend*… à part de d'ça je vois pas d'autre lien… (TR/MA)

(31) Ey... j'en ai aucune idée de cette mot là… euh, peut-être ben comme *knee* en anglais, c'est comme un genou… mais cette mot là, c'est *Knie* fait que… non, ça me dit rien, j'en ai aucune idée… (TR/N)

(32) *Wetter*… euh… je peux vraiment pas dire c'est quoi… c'est tannant je sais pas, j'en ai aucune idée, ça me dit rien pantoute, pantoute, pantoute, je peux pas dire à quel mot ça pourrait ressembler... je vois aucune ressemblance… peut-être que ça ressemble à un autre mot en allemand qui se dit presque pareil mais j'en ai aucune idée… (nonTR/A)

In the excerpt above, the analysis of *Freund* reflects a cross-check with the sentence thus confirming the initial association with the English cognate *friend*. In the following example, the participant recognized the similarity between *Knie* [kniː] and *knee* [niː], but did not attempt to analyze the sentence at all. Instead, uncertainty about the choice was stressed by underlining that *Knie* was different, whereas the similarity relation between *Freund* [ˈfrɔynt] and *friend* [ˈfrend] could have raised more serious doubts. In fact, the example (31) appears to reflect a lack of confidence. Finally, when faced with the word *Wetter* [ˈvɛtɐ], it really seems to be self-consciousness that prevented the participant from

162

attempting any translation at all. Essentially, these observations point to the fact that depth of analysis was not consistent at the individual level. Moreover, they suggest that personal and contextual factors might influence the kind of reasoning that was displayed for a given item at a given moment. For instance, the shallow analysis of the items exemplified under (31) and (32) could stem from a build-up of self-doubt as the participant was working through the task. On the other hand, analyses that were coded MA or UR (see examples 21 through 29) did not include any markers of uncertainty. These observations suggest that depth of analysis might also be related to individual levels of confidence. On the other hand, it is also possible that the uncertainty expressed in (32) was simply due to the absence of an initial association. That is, when no possible answer comes to their mind right at the start, some learners seem to give up trying to find one. Whether this is due to self-consciousness or merely to a lack of interest, cannot be inferred from the data. Furthermore, the degrees of similarity between the target and source items as well as the learner's degree of familiarity with the English cognate might be at play here.

In sum, the observed strategies and the respective levels of awareness that were identified in the TAPs should not be taken as a global measure of the actual abilities of each participant to reflect upon and manipulate the target items and structures. In turn, the variability that was observed both across the sample and at the individual level points to the complexity of the concept of MLA, the manifestations of which are possibly interrelated with other individual and contextual factors, especially when the measure of MLA implies on-line processing, i.e., the active application of metalinguistic knowledge in a given learning situation.

The overall results of the qualitative analysis of the TAPs support our hypothesis concerning the positive relationship between metalinguistic awareness and positive transfer: Even if it was possible to transfer on very low levels of awareness and if higher levels of awareness did not necessarily lead to transfer, those participants who transferred a lot also tended to analyze word and sentence structures more thoroughly. Namely, parametric correlations between the raw scores of the transfer measure and the raw MLA scores from the TAPs proved to be highly significant ($r = .57$, $p < .001$). An overview of

the relative distribution of the levels of MLA exhibited in the TAPs in relation to the measure of positive transfer is provided in Figure 15.
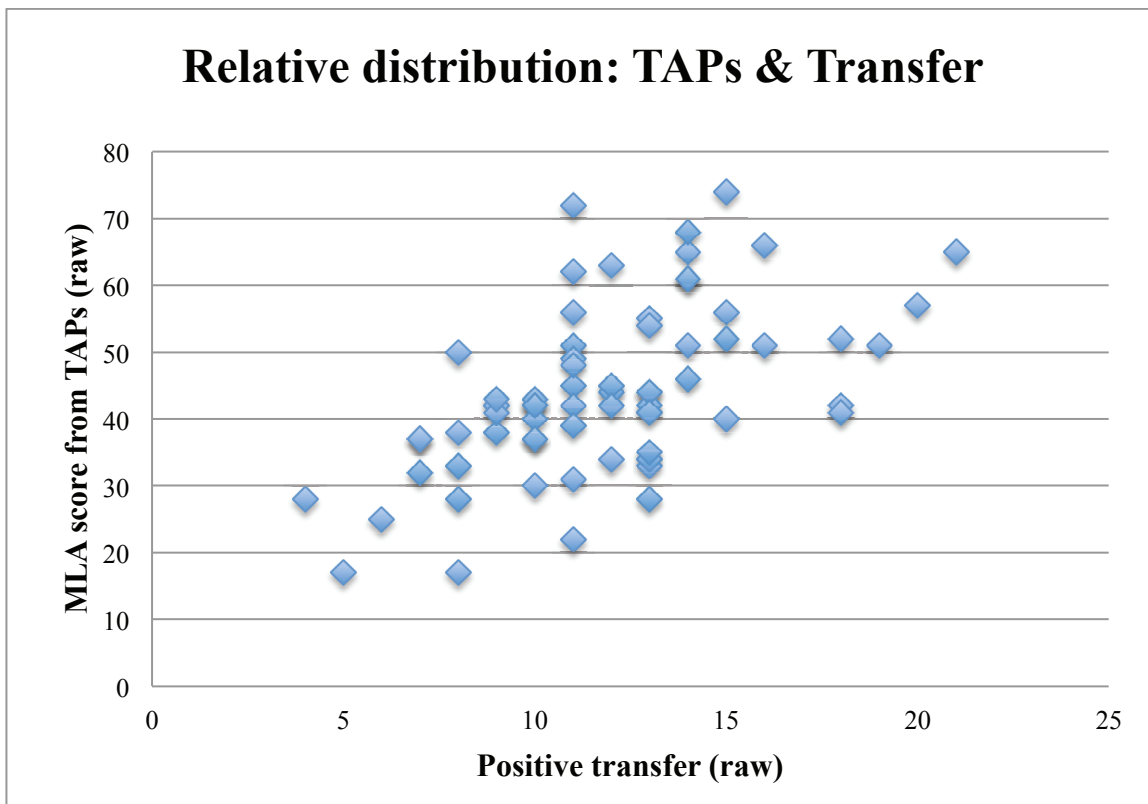
Figure 15. *Relative distribution (MLA score from TAPs / Positive transfer)*
Note. This figure represents the raw scores obtained on both measures.

This relationship will be further discussed in the last part of this chapter when elaborating on the regression model that was used to determine which of the different influential factors appeared to be the strongest predictor of positive transfer from English (see section 5.5). However, it should already be pointed out that this particular measure of MLA was based on the same data as the transfer rates. As previously mentioned, the verbalizations that coded either *Absence* or *Prior knowledge* were both rated zero (0). At the same time, both these levels also reflected absence of transfer, which was rated zero (0) as well. Hence, even if the actual levels of metalinguistic analysis I, N, MA or UR, respectively, did not depend on whether or not transfer had taken place, the fact that a part of the transfer measure was somehow included in the coding of the TAPs, namely at level zero, the relationship between the two measures was at least partly predictable. Note however that

when both measures were normalized with respect to instances of prior knowledge in the transfer measure and zero-ratings in the MLA score from the TAPs, they were also strongly correlated ($r = .58$, $p < .001$). This conceptual and methodological aspect of the relationship between MLA and positive transfer will be further discussed in the following chapter, as we procede to answering our central research question in section 6.1.2.

5.2.2.3 Inter-rater reliability

As previously mentioned, we also had two independent raters score the think-aloud protocols for levels of metalinguistic analysis. This was particularly important given that the categories for this coding procedure had been established by the main researcher herself. Across the sample, we listed 142 deviations in coding between the two raters, which represented 7.68 % of the 1848 verbalizations. The categories that were most frequently interchanged were those of Noticing and Meta-Awareness, which happened precisely 55 times. This was mainly the case for verbalizations where the resemblance to a similar word was explicitly stated, including a reflection on the accompanying sentence. However, whereas one of the raters would count this reflection as an explicit cross-check with co-occurring words and/or structures and thus rate the verbalizations on the level of MA, the other rater would not take this cross-check as a justification of the respective answer and thus use the code N. The examples listed below illustrate this for three different test items.

(33) Avec la phrase ouais, ouais, ouais, ça serait des des bas qu'on perd toujours dans la laveuse, puis en dedans … a *sock* [sak] a *sock* [sok], les bas les les les chaussettes qu'on perd toujours dans la sécheuse… et qu'on a jamais une paire, qui est dur de trouver… (ID 10: MA/N)

(34) « Morgen gibt es einen Sturm »… morgen, c'est le matin, peut-être… « ce matin il y avait une »… ça me fait penser à une *storm*… de l'anglais, fait que ça doit être un éclair… (ID 49: N/MA)

(35) Schiff… Schiff… « ein Schiff fährt über den Fluss »… *Schiff*… eum… *Schiff*, ça serait-tu comme *ship,* un bateau ? (ID 5: MA/N)

From these examples, it follows that the descriptors for each category might have to be slightly more detailed. In this case, for instance, it would have to be indicated whether any

kind of explicit sentence check qualifies the verbalization as metalinguistic. Precisely, in the examples listed under (33) and (35), the reference to co-occurring words in the sentence appears as a confirmation of the translation choice, whereas the reflection on the sentence under (34) may or may not have played a role in the activation of the English cognate *storm*.

Another distinction that caused some deviation was that between the levels of MA and I in verbalizations that might merely appear as random clang associations at first sight, but where the participants analyzed forms and meanings of other words, resulting from a cross-check with the sentence or with other learning events and which were thus rated at the level of MA, if that analysis appeared logical to the rater. The following two examples illustrate this difference in interpretation for each of the raters.

(36) *Wetter*… ça ressemble à *water* en anglais, mais c'est pas ça parce que en allemand ben eau c'est *Wasser*…*Wetter Wetter*… [prononce la phrase] *Wetter, Wetter, Wetter*... ça ressemble à *better*, mais non… (ID 22: MA/I)

(37) *Schiff*… euh… ça me fait vraiment penser à *chiffre* en français là… euh, je pense pas que c'est ça… je pense pas que ça a non plus un quelconque lien avec les nombres… ça me fait penser à *shift* en anglais… euh, mais c'est un verbe… peut-être que ça pourrait être… (ID 37: I/MA)

In the example under (36), the reference to the German word *Wasser*, based on the English cognate *water* made it clear to the participant that the target item *Wetter* could not mean *water* as well. When comparing it to *better*, this choice was also rejected. The first rater judged this line of thought as logical reasoning based on other words in both English and German, and thus rated the verbalization at the level of Meta-Awareness, whereas the second rater judged the abundance of alternative answers as random clang associations, rated at the level of Intuition. In the example under (37) however, it was the first rater who dismissed the participants' reflections on the target item Schiff as merely intuitive references to similar-sounding words from other languages, whereas the second rater qualified this verbalization as MA. A possible explanation for this rating could be the explicit reference to word class, which led the participant to reject *shift* as a possible translation. However, given that *shift* could also be a noun, this logic may or may not be

166

interpreted as a sufficient justification. As for the distinction between the levels of Noticing and Meta-Awareness, it might have to be clarified in the coding procedures what types of reasoning qualify as "justifications based on internal logic" (see Appendix E).

Moreover, it should be kept in mind that insufficient knowledge of German might also be the cause for some isolated inconsistencies in rating, which might not have occurred if the second rater had also had native-like understanding of German.[57] In this regard, the following two examples illustrate cases where the analysis of some co-occurring words in the sentence revealed a total absence of internal logic, whereas the second rater, who did not have sufficient understanding of the German sentences, rated the verbalizations at the levels of MA and I, respectively.

(38) « Martina hat ein gutes Herz »… *Herz*… ça doit être des… euh, pff… des goûts, je dirais des goûts… puis *Herz* ça doit être spéciaux… j'ai aucune idée… *goût* à cause de *gutes* là… (ID 49: A/MA)

(39) *Haar*… [marmonne fin de la phrase]… *lockiges*, ça me fait penser à un *locker,* une case… *Haar*, ça doit être un cadenas, sinon ça me fait penser à aucun autre mot… j'ai aucune idée c'est quoi… ça doit être un cadenas… (ID 49: A/I)

In the former example (38), the participant starts off by translating the target item *Herz* as *goûts*, the plural of *taste* in French, based on the adjective *gutes* (English: *good*) that co-occurs with *Herz*. Then, she goes on to translating *Herz* as *spéciaux* (English: *special*), an adjective that appears to co-occur frequently with *goûts* in French. This verbalization clearly reflects a lack of grammatical awareness, since the notion of word class does not appear to have any impact on the learner's reasoning. The same confusion regarding the word class of the target item in relation to co-occurring words in the sentence is displayed in the latter example (39), where the same participant associates the adjective *lockiges* (English: *curly*), which precedes the target item *Haar* (English: *hair*), with the English word *locker* and thus translates *Haar* as *cadenas*, which is the French translation equivalent to *padlock*. Again, the associations were not only random, but also totally illogical.

---

[57] We acknowledge that it would have been ideal to have a second rater with sufficient knowledge of German, who would thus have easily detected misinterpretations of the target sentences as the ones exemplified under (38) and (39), respectively.

However, it was necessary to understand the sentence that was displayed with each of these test items to grasp the extent to which the answers were devoid of internal logic.

Finally, in order to provide a critical analysis of the coding procedures established by the main researcher, one other distinction between categories that appeared to cause some deviation should be mentioned here. In fact, it was not always clear what kinds of verbalizations should be rated at the highest level of metalinguistic analysis, namely UR, as opposed to MA. As previously mentioned with respect to this last category, the task type did not necessitate rule formulation. Hence, the category UR was also supposed to cover cases where a thorough analysis of co-occurring words or other related meanings were taken to justify the translation. However, the two examples below illustrate that the line between the levels is not an easy one to draw.

(40) *Sonne…* ça ressemble à, au mot *Sohn*, ben, un fils en allemand, sauf que il y a pas le *h* puis c'est pas au pluriel non plus euh, « Die Sonne geht auf »… hmm, ça pourrait sonner comme *Sohn*, mais non, c'est pas *Sohn*…c'est *Sonne, Sonne*, [prononce la phrase de nouveau] (ID 22: MA/UR)

(41) *Freund*, c'est un ami… ou un chum… un amoureux parce que je le sais parce que on l'a appris en allemand, mais sinon je l'aurais su parce que ça ressemble à *friend,* mais je suis tout le temps mêlée entre *Freud* et *Freund* je trouve… alors je sais pas c'est quoi Freud, mais j'aurais pu me mêler puis dire « das ist mein Freud » puis ça n'aurait vraiment pas voulu dire la même chose mais en tout cas… (ID 58: UR/MA)

Both the above verbalizations display a thorough analysis of the target item in relation to other words in German that could have led to confusion but that are explicitly rejected on the basis of slight differences in sound or spelling. In fact, both examples could be rated at the highest level of awareness given that both learners display a conscious control over the vocabulary that they have learned, as they analyze the formal differences between similar items of the target language appropriately. What could have influenced the divergent ratings regarding the respective degree of analysis was the fact that the participant in the former example did not find the appropriate translation of the target item, which may have been interpreted as less analyzed knowledge. At the same time, the type of analysis

illustrated in the latter example could also be judged as requiring less depth, given that this learner already knew the target item prior to consciously manipulating its form and meaning. At the same time, the statement was formulated as a sort of reminder for subsequent learning situations. If interpreted as a rule of thumb to prevent confusion based on the same word pair in the future, this level of abstraction might in fact be taken to equal depth of analysis.

In sum, what should be retained from these observations is that the ratings pertaining to depth of metalinguistic analysis are to some extent influenced by the basic content as well as the appropriateness of the answer itself. By extension, this naturally encourages a preference for a measure of MLA that does not have the potential to interfere with another measure, in this case that of transfer from English. However, it must be acknowledged that the categories of analysis ranging from I to UR were shown to be largely independent of the transfer measure, as discussed in the previous sections. Furthermore, the categories appeared to be sufficiently specific for two independent raters to evaluate the sample with a high degree of concordance. On the whole, inter-rater reliability was of 92.3 %. Finally and most importantly, it proved to be very useful to include the metalinguistic analysis of the TAPs as a secondary measure of MLA in order to shed light on the different kinds of cross-linguistic processing that were displayed throughout the sample. As will be exposed and discussed in the following subsection, the ability to reflect upon and manipulate one's own native language appears to provide a sound basis for the conscious activation of potentially helpful correspondences across the different languages known to the learner. In other words, the type of MLA that was assessed in the THAM-3 was positively correlated with the metalinguistic measure of the TAPs. However, each of the tasks appeared to require different types of processing, which were largely context-related when it comes to the application of metalinguistic abilities to a particular multilingual learning context, as revealed by the degree of inter- and intra-individual variability observed in the TAPs.

5.2.3   Correlations between measures of MLA

In line with the conception of multilingual proficiency conveyed by the DMM (Herdina & Jessner, 2002), the types of awareness assessed in the present study were considered as two

sides of the same coin. Whereas the performance on the THAM-3 was taken to represent the ability to analyze and manipulate different types of linguistic structures of participants' native language French, the verbalizations extracted by means of the TAPs revealed the active application of metalinguistic abilities in an exercise that involved the manipulation of mostly unknown vocabulary in their L3 German. The kind of awareness exhibited in the THAM implied a reflection on the structural features of their own language, as they know it, and was thus referred to as "reflexive" (see section 4.5.1.2 on coding procedures). In contrast, the kind of awareness exhibited in the TAPs implied an active search across their multilingual lexicon during which metalinguistic abilities could be "applied" to confirm specific translation choices. On the one hand, higher levels of reflexive awareness were expected to provide a basis for the capacity to apply analytic skills to a multilingual context and thus reveal higher levels of applied crosslinguistic awareness. In other words, these complementary measures of MLA were expected to be positively correlated across the sample. However, several sources of variability, which appeared to be mostly context and task-related (see sections 5.2.1 and 5.2.2), must be taken into consideration to interpret the correlations between the two measures, which were moderate but significant, as illustrated in Table 16 below.

Table 16. *Parametric Correlations (TAP / THAM)*

		MLA score from TAPs (raw)	MLA score from THAM
MLA score from TAPs (raw)	Pearson Correlation	1	.31[*]
	Sig. (2-tailed)		.01
	N	66	66
MLA score from THAM	Pearson Correlation	.31[*]	1
	Sig. (2-tailed)	.01	
	N	66	66

*Note. * $p < 0.05$.*

As previously stated, performance on the THAM-3 seemed to be influenced by a certain lack of motivation or interest on the acceptability part of the test, which could make a notable difference in the total score given the relatively large number of test items in this section (N = 12). For instance, a closer look at the top left corner of the scatter plot below (Figure 16) reveals that three participants reached very high scores on the MLA measure of

170

the TAPs whereas their scores on the THAM were particularly low (ID 54: 72/6, ID 56: 66/9, ID 1: 65/10). Not surprisingly, the low THAM scores of these three participants were clearly related to the fact that the text had not been analyzed at all (ID 56) or only partially and at the surface level. In fact, none of the two participants (ID 1, ID 54) who did attempt to correct a few errors and justify their reasoning reached even an intermediate level of ML = 1 for any of the items, whereas all three of them demonstrated MLA in their analysis of the first four items, both at level 1 and 2. Thus, even if they failed to convince the raters regarding their ability to reflect on metasyntactic and grammatical structures, their capacity to reflect on other types of structural features of language could be confirmed in their high MLA scores from the TAPs.
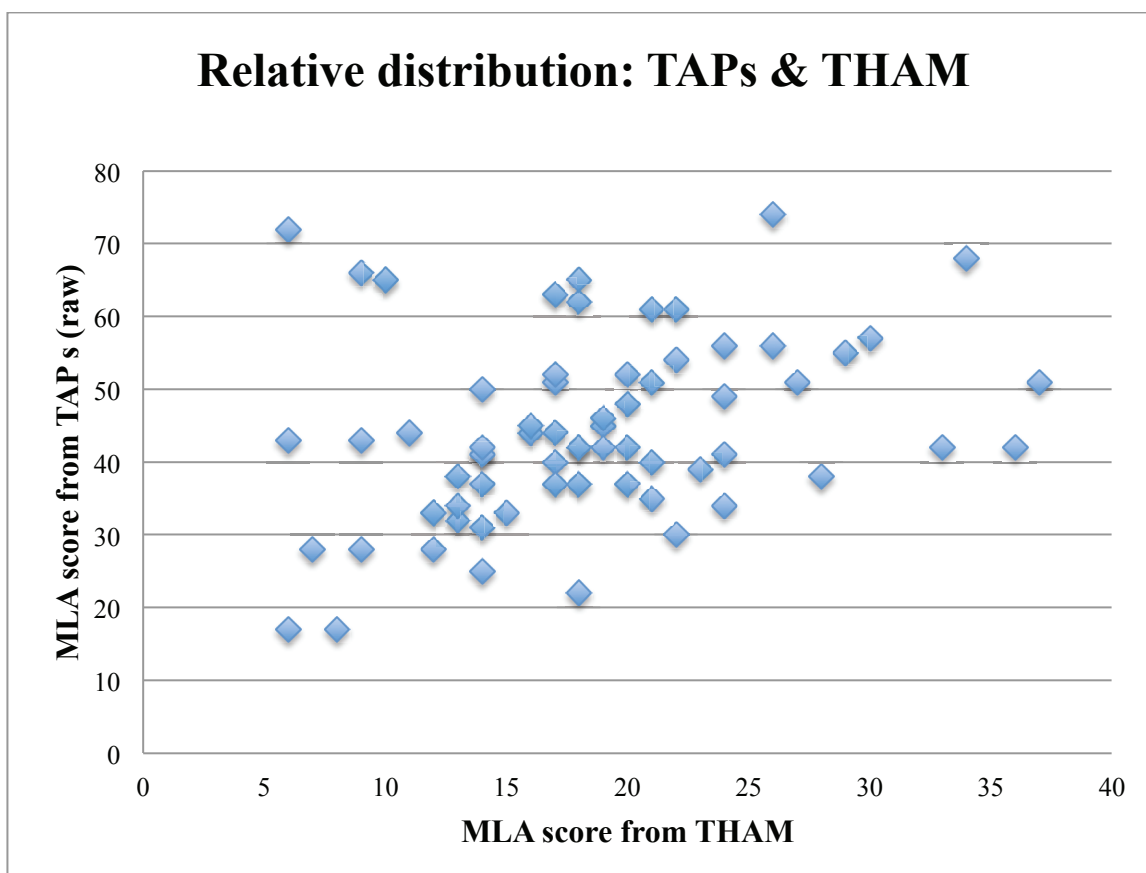
Figure 16. *Relative distribution (MLA scores from TAPs / MLA scores from THAM)*

In addition to these few outliers in the top left corner of the graph, it must also be acknowledged that certain participants appeared to display nearly the opposite distribution, namely particularly high scores on the THAM in combination with only medium sized

scores on the TAPs, as illustrated by the group of three in the mid-right section of the graph (ID 15: 42/33, ID 39: 42/36, ID 61: 51/37). Two of these participants (ID 39, 61) had demonstrated a particularly meticulous analysis of the majority of the items of the THAM, whereas their verbalizations during the translation task were mostly based on obvious similarity relations, thus revealing a level of *Noticing* for the majority of the items of the translation task. Moreover, their reflections were concise and never more elaborate than necessary to infer a translation equivalent. Also, both of them simply stated that they did not know the answer in six (6) of the cases (coded as *Absence* = 0), instead of trying to infer possible related meanings by analyzing the sentence for contextual cues. These examples illustrate to what extent levels and types of metalinguistic analysis may be related to personal preferences for verbalizing thoughts. That is to say that the careful performance on the THAM, which revealed high levels of MLA for these participants, was not actually in contrast with their concise answers in the translation task. Rather, they simply solved the task as they were asked and reached top scores in the transfer measure (ID 39: TR = 18, ID 61: TR = 19) without having to resort to deeper levels of metalinguistic analysis to do so. While this observation points to distinct types of reasoning across individuals, it also calls attention to the limitations of the translation task in itself, in the sense that it was possible to translate on the basis of rather simple associations. In other words, the extent to which MLA was displayed in the TAPs did not necessarily reflect the degree of MLA that participants might actually be capable of when working on other types of tasks. From a methodological perspective, this critical analysis underscores the need for further examination of this variable. Moreover, it emphasizes the importance of a careful interpretation of qualitative data that has been quantified for comparison purposes. The kinds of reasoning reflected in the verbalizations of the third participant mentioned above (ID 15) were for the most part random clang associations (coded *Intuition* = 1), which did not lead to positive transfer from English (Total TR = 9). However, six (6) items were analyzed at the level of MA (coded 3), which added up to a MLA score comparable to that of ID 39 and ID 61 presented above, except that their approaches to the task were strikingly different. Namely, this third participant (ID 15) verbalized every possible association, both with similar sounding words and with contextual cues, but he/she did not question his/her own illogical statements. This way of approaching the task led to considerably less positive

172

transfer, but still reflected the capacity of analyzing co-occurring linguistic resources correctly when possible distractions from other clang associations were limited, namely in the cases of *Herz, Milch, Maus, Socke, Knie* and *Stein*, which were coded at the level of MA and successfully transferred in four (4) cases.

In sum, the two measures of MLA were clearly related to each other with regard to the very core of the notion itself, i.e., *deliberately focused attention on the structural features of language*. However, the types and levels of MLA exhibited in the THAM and in the TAPs, respectively, could reflect different kinds of reasoning at the individual level or differences in verbalization of the reasoning. Inter- and intra-individual variability regarding levels of MLA across these two measures could be attributable to affective factors such as motivation or self-consciousness, or to individual context-related factors such as boredom, distraction or fatigue and finally to the type of task itself.

5.3   English L2 measures: proficiency and exposure

This section will be devoted to the descriptive results of the measures that are related to the participants L2, English. The predictive variables that were examined in relation to positive transfer were proficiency in and exposure to the English language. Proficiency levels were determined by means of the Michigan Test of English Language Proficiency (MTELP), whereas relative exposure to English was assessed by means of a background questionnaire. The following subsections provide an overview of the results obtained on these L2 measures, respectively.

5.3.1   English language proficiency according to MTELP scores

As outlined in Chapter 4, the MTELP consisted of three sections for which different partial scores could be attained: Grammar (40 points), Vocabulary (40 points) and Reading comprehension (20 points). Table 17 provides an overview of the descriptive test results for each section separately, for the total raw score and for the equated score. Namely, due to the level of difficulty of the version of the test that was used, raw test scores were equated on the basis of a reference table provided by the test authors (Corrigan et al., 1979). If the goal of this standardized test had been to assess whether these learners' proficiency levels

were sufficient to access an English language university, the appropriate measure to use would be the equated score. Likewise, one would refer to equated scores if participants had taken different versions of the test. However, given that a single version of the MTELP was used to compare learners of a single group among each other, the variable of English language proficiency was operationalized based on the raw scores.

Table 17. *Descriptive Statistics (MTELP partial and total scores)*

	Min	Max	Mean	SD
Grammar (40)	11	40	28.94	7.37
Vocabulary (40)	8	40	26.73	7.02
Reading Comprehension (20)	4	20	12.53	3.76
MTELP total (raw) (100)	28	99	68.09	16.81
MTELP total (equated) (100)	44	99	79.39	11.14

Interestingly, from comparing the means of each separate section to that of the total score, it appears that participants who performed rather weakly on one part of the test did not do so across the board, but that one or two sections may have caused them more trouble than others. The percentages of the partial mean scores suggest that reading comprehension was more difficult for our participants than grammar (Grammar: 72.3 %; Vocabulary: 66.8 %; Reading: 62.7 %; Total: 68.1 %). To illustrate the relative distribution of the partial scores, two figures were included, in which the performance on the reading section (x-axis) was paired with the total grammar and vocabulary scores, respectively (Figures 17 and 18).
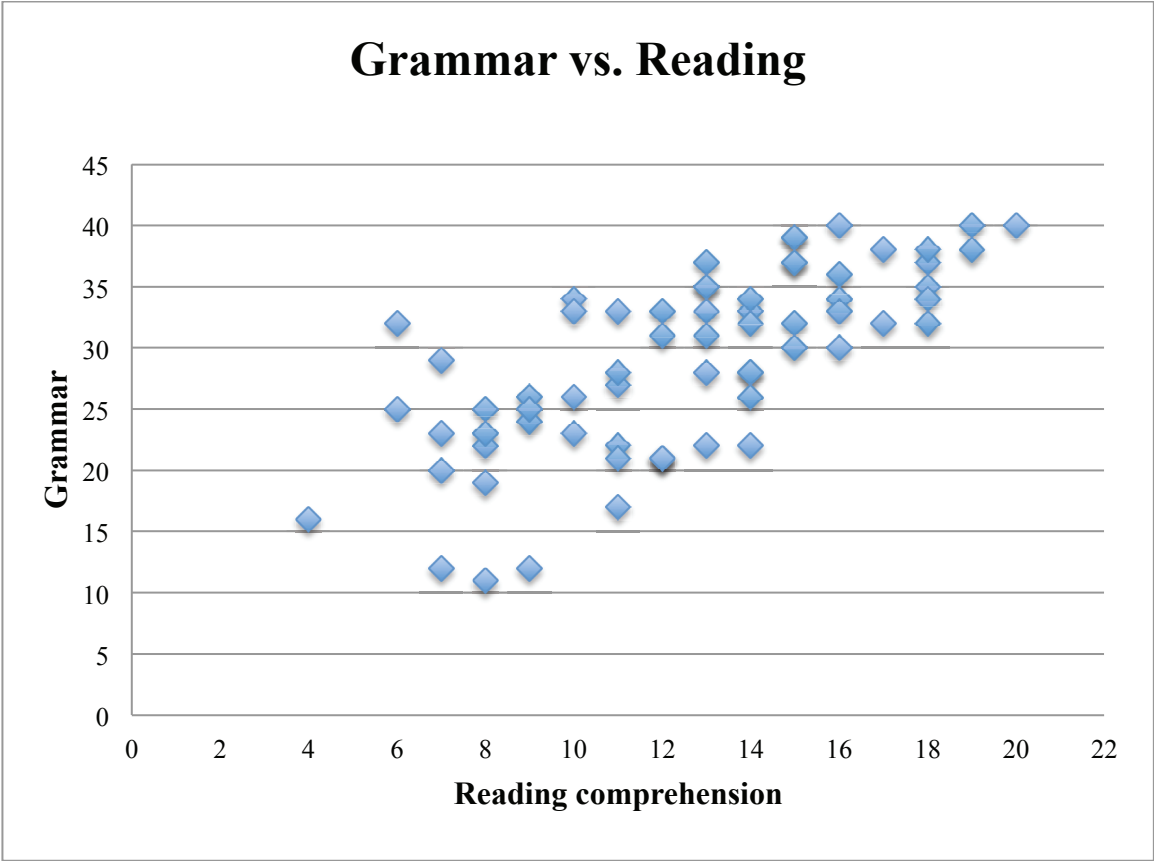
Figure 17. *Relative distribution (Grammar / Reading comprehension)*
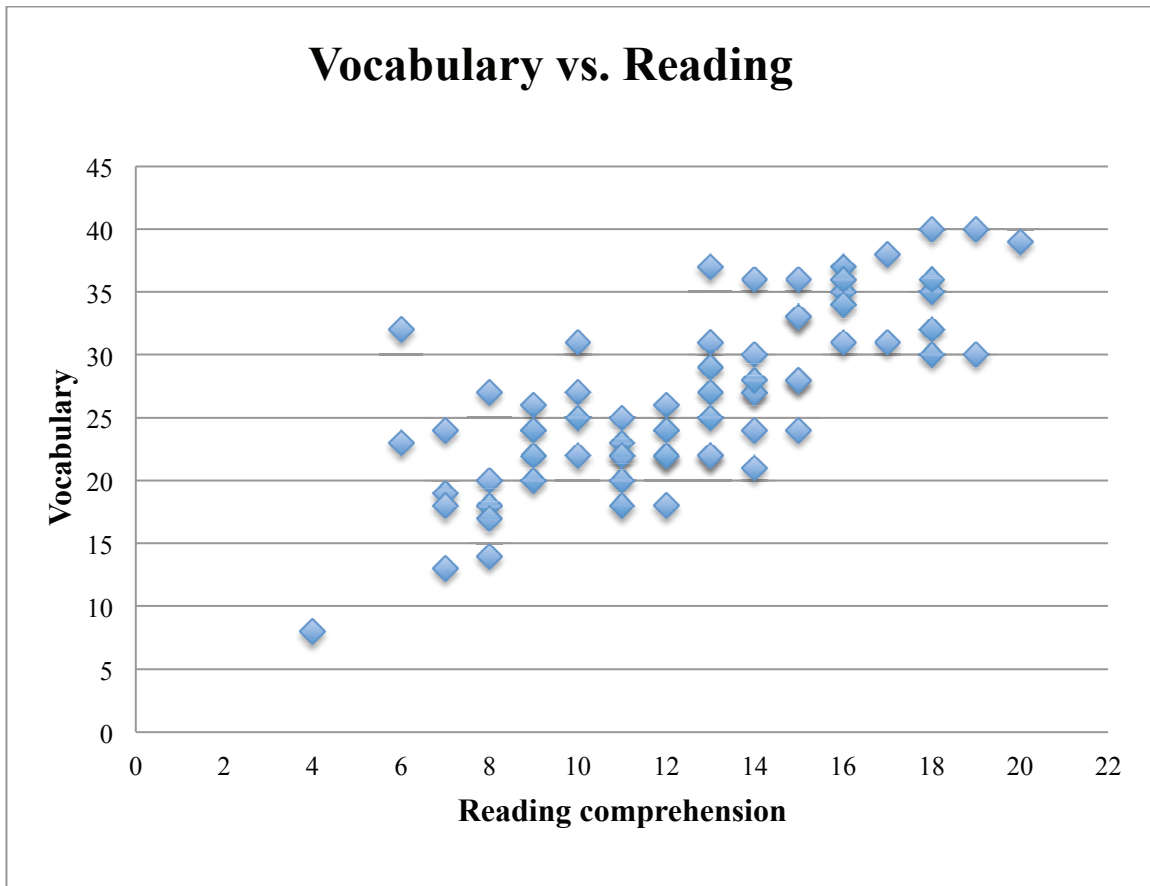
Figure 18. *Relative distribution (Vocabulary / Reading comprehension)*

Even if the mean score on the reading part was slightly lower in relation to the two other proficiency measures, a close look at the two figures above reveals that the indivudal performance on the test was rather consistent across sections. That is to say, weak reading performance was most often associated with relatively weak grammar and vocabulary knowledge, while the strong readers of our sample also performed well on both grammar and vocabulary measures. Still, both figures reflect a tendency towards weaker reading performance in relation to either of the other measures, as illustrated by the participants who appear in the top left corner of each graph. For instance, seven participants who only received between 6 and 8 points in the reading section, reached a grammar score between 22 and 32, and four of them also performed well above average on the vocabulary section, thus suggesting that reading was in fact what caused our partcipants most trouble in this proficiency test. However, on the whole, the results suggest that the level of difficulty was appropriate to measure proficiency levels of the targeted population. Namely, maximal

scores were reached in each section while means and standard deviations reflected a rather large range of results across the sample (see Table 17).

5.3.2 Exposure and frequency of use

As outlined in Chapter 4, different kinds of exposure data were collected by means of a background questionnaire. Precisely, a rather basic type of exposure to English was addressed in Q8 of the questionnaire, in which participants were asked to report on the time spent in different linguistic environments. However, the primary measure of exposure in the present study consisted of a grid presented in Q21, in which participants were asked to rate the relative frequency of use of each of their languages in ten distinct contexts of use. The following subsection provides the results obtained after quantification of the relative frequencies of use, after which the secondary exposure measure will be discussed.

5.3.2.1 Quantification: Frequency of use

Unfortunately, it was beyond the scope of the present study to keep track of the activities that our participants truly engaged in over a certain period of time, such as by means of a diary or an activity log (e.g., Ranta & Meckelborg, 2013). However, the ten contexts of use exposed in the grid adapted from Grosjean, Jaccard and Cividin (2000) included receptive, productive and interactive activities, in both written and spoken mode. This allowed for an analysis of distinct types of language use in relation to L2 proficiency and possibly even to positive transfer, thus reflecting a slightly broader conception of exposure than reflected by traditional questionnaire-type investigations into the amount of contact with the target language. As previously stated, frequency ratings ranged from never (0) to several times per year (1), per month (2), per week (3) and daily (4). Table 18 provides an overview of the descriptive results regarding the reported frequency of English use for each activity, in addition to the total frequency of use (sum of frequencies) listed in the last line.

Table 18. *Descriptive Statistics (Frequency of use across activities)*

	Min	Max	Mean	SD
interagir oralement: amis (4)	0	4	2.41	1.15
interagir oralement: famille (4)	0	4	1.45	1.35
écouter radio (4)	0	4	1.74	1.49
lire: journaux, revues (4)	0	4	2.11	1.39
lire: livres (4)	0	4	2.45	1.26
écrire: lettres, courriels (personnels) (4)	0	4	2.00	1.32
écrire: lettres, courriels (formels) (4)	0	4	1.27	1.27
écrire: textos, clavardage (4)	0	4	2.86	1.15
regarder: émissions, films, series (4)	1	4	3.15	.92
jouer: jeux video (4)	0	4	1.95	1.56
Frequency of use: English (40)	7	40	21.44	7.40

Note. 0 = 'never'; 1 = 'several times per year'; 2 = 'several times per month'; 3 = 'several times per week'; 4 = 'daily'

The activities for which the most frequent use of English was reported were 'watch television or films' (Mean = 3.2), 'write text messages/chatting on the internet' (Mean = 2.9), 'read books' (Mean = 2.5) and 'talk to friends' (Mean = 2.4), to name just a few of them. However, daily use of English was listed for each of the listed activities by at least 5 participants, namely for 'writing formal letters or emails', and at the most by 31 of them, namely for 'watching TV or films'. There was no notable tendency across the sample as to preferences for receptive, productive or interactive activities in English. However, a principal component analysis revealed distinct loadings for three groups of activities. Table 19 provides an overview of the respective factor loadings.

178

Table 19. *Rotated Component Matrix[a] (Frequency of use across activities)*

	Component		
	1	2	3
interagir oralement: amis	.31	**.67**	.11
interagir oralement: famille	-.12	**.86**	.28
écouter radio	.23	**.69**	-.30
lire: journaux, revues	**.77**	.18	-.26
lire: livres	**.60**	.16	.38
écrire: lettres, courriels (personnels)	**.75**	.05	.33
écrire: lettres, courriels (formels)	**.67**	.21	.10
écrire: textos, clavardage	**.67**	-.02	.34
regarder: émissions, films, séries	.38	.18	**.68**
jouer: jeux vidéo	.09	-.02	**.81**

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 6 iterations.

The first component that could be clearly distinguished from the two others included the two reading and the three writing activities and was therefore labeled 'written interaction and reading'. The second component included oral interaction with friends and family as well as listening to the radio, which explains the label 'oral interaction and listening'. Finally, the two activities that involved exposure to a screen, either of a television or of a computer, were also grouped together under the label of 'interaction with a machine'. What these loadings indicate is that those participants who used their English more frequently for one type of oral interaction were also likely to do so for other oral activities. Likewise, the frequent use of English in one writing-based activity was also likely to apply for the other activities in this mode of communication, no matter whether the type of writing was formal or very informal, such as, for instance, in text messages or chats. Finally, those who watched a lot of TV shows or films in English also tended to play video games in that language. Based on these factor loadings, the primary exposure variable labeled 'Frequency of use' was divided into three subcategories, in order to examine the relationship of each of these types of exposure and use to overall proficiency in English, as measured by means of the MTELP.

5.3.2.2 Quantification: Basic exposure

The secondary measure of exposure to English was based on a question (Q8) in which participants were asked to list the number of weeks spent four different linguistic environments: a country or a province, a family, a work place or a school. In case two or more indications referred to the same experience, such as a work place in an Anglophone province, for instance, they were asked to indicate the number of weeks for only one of these environments. Furthermore, separate shorter stays in any of these environments were supposed to be added up to a total number of weeks, or months, if applicable. Not surprisingly, some participants had never traveled whereas others had spent time in different English-speaking contexts repeatedly throughout their lives, which resulted in abnormal data. That is, total numbers of weeks above 100 and even up to 316 were listed for only three participants, whereas the majority of them had spent very little or no time at all in any of the listed environments, as illustrated in the dot plot in Figure 19.
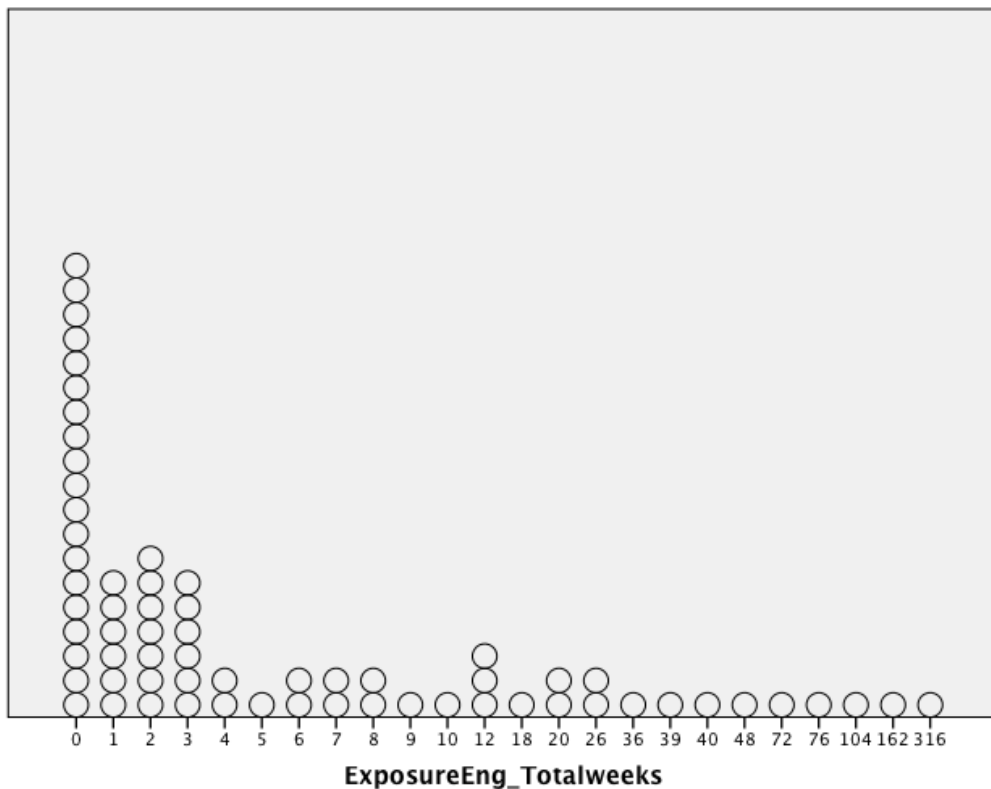
Figure 19. *Participants per exposure level: Number of weeks in English-speaking environments*

Nevertheless, the aim of triangulating data based on different exposure measures motivated further analyses. Precisely, if the amount of time spent in English-speaking environments was to be investigated as a possible predictive variable, along with 'Frequency of use', it had to be possible to establish some sort of relationship between these two exposure measures. Namely, if the results obtained on Q8 could reliably represent the amount of time spent in the listed environments where English was used, then the amount of time that the participants themselves spent using their English for different language-based activities should be somehow related, because opportunities of using that language are multiplied in an English-speaking environment.

5.3.2.3 Relationship between exposure measures

In order to examine the relationship between the two exposure measures, we ran a series of parametric correlations between the number of weeks spent in English-speaking environments and the total frequency of English use added up from the partial frequencies reported in Q21. At first, basic exposure was quantified by adding up the total number of weeks reported in Q8. Given the abnormal distribution of the total number of weeks across the sample, as discussed in the previous section, it was not particularly surprising that no correlation could be established between the two variables (Pearson's $r = -.04$, $p = .73$). Given the presence of extreme values, we also ran non-parametric correlations to validate the fact that no relationship could be established between the two measures, even when outliers did not affect the equation. Table 20 below illustrates that Spearman's correlation coefficient was also negligible.

Table 20. *Non-parametric Correlations (Exposure to English / Frequency of use)*

			Exposure to English (Total: weeks)	Frequency of use: English
Spearman's rho	Exposure to English	Correlation Coefficient	1	.10
		Sig. (2-tailed)		.43
		N	66	66
	Frequency of use: English	Correlation Coefficient	.10	1
		Sig. (2-tailed)	.43	
		N	66	66

In order to examine the possibility that this absence of a relationship between the two exposure measures was due to the wide range of results represented in the total number of weeks, the variable was divided into five larger units, precisely 'never', 'one week or more', 'four weeks or more', 'three months or more' and '1 year or more', rated on a scale from 0 to 4. When these categories were applied to the total number of weeks, there was still no significant correlation to be established between exposure and frequency of use (Pearson's r = .13, p = .30 / Spearman's *rho* = .10, p = .42). However, a quick look at the scatter plot below (Figure 20) reveals the difficulty of comparing categorical with continuous variables.
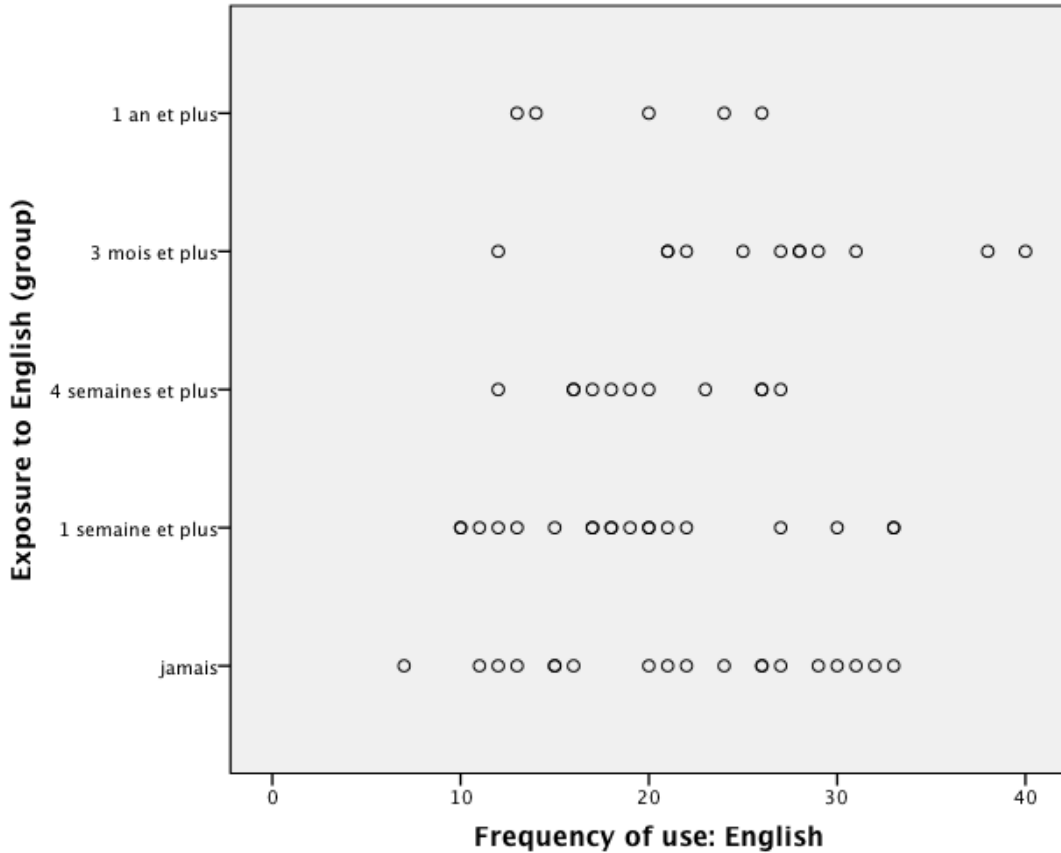
Figure 20. *Relative distribution (Exposure per group / Frequency of use)*

Finally, it seemed most reasonable to transform 'basic exposure' into a continuous variable, based on the same quantification pattern applied to 'Frequency of use'. That is, the five categorical units of frequency ('never', 'one week or more', etc.) were quantified for each of the four language environments according to a five-point scale ranging from zero (0 = 'never') to four (4 = 'one year or more'), in accordance with the scale established for the relative frequencies of use ranging from 'never' (0) to 'daily' (4). At the end, these quantified levels of frequency for each environment were added up to form a total frequency count, with a possible maximum of 16. Descriptive results of the variable thus labeled 'Exposure to English', including its subcomponents, are listed in Table 21.

Table 21. *Descriptive Statistics (Exposure to English across environments)*

	Min	Max	Mean	SD
WeeksEngCountry (4)	0	4	1.23	1.09
WeeksEngFamily (4)	0	4	.50	.10
WeeksEngWork (4)	0	4	.35	.94
WeeksEngSchool (4)	0	4	.30	.84
Exposure to English (12)	0	12	2.38	2.78

Note. 0 = 'never'; 1 = 'one week or more'; 2 = 'one month or more'; 3 = 'three months or more'; 4 = 'a year or more'

Ultimately, a statistical relationship could be established between this basic measure of exposure to English and the frequency of English use. Even if the correlation was only weak to moderate, it was significant (Pearson's $r = .27$, $p = .03$).

Of course, statistical results involving continuous variables that are derived from categorical ratings must be interpreted with care. Even if variable formation for both exposure measures was comparable, it must also be born in mind that the respective question types required different kinds of information, which might in fact put into question the reliability of the secondary exposure measure. That is, the reported number of weeks spent in the respective linguistic environments was to be added up over an average lifetime of 18 years. The accuracy of these data therefore depended on the ability to recall, the ability to add up exposure times and on the willingness to do so. Moreover, it was possible that the categories of linguistic environments were understood differently. For instance, participants working in the service sector in the province of Québec may have counted periods of time when they served Anglophone clients in English as time spent in a English-speaking work place, whereas others may have interpreted this category in terms of a totally Anglophone environment. Likewise, what may have been included in the family category ('une famille où l'on parle l'anglais') were prolonged periods of interaction in English with members of their own or another non-Anglophone family. This assumption was based on the fact that a number of participants who later confirmed to have grown up in monolingual families indicated in Q5 of the questionnaire that they spoke other languages (including English) with certain family members. Given that subjects who had actually grown up in bi- or multilingual families had to be excluded from the sample, these

cases were detected directly after the first analysis of the questionnaire. By contrast, no specification questions were asked concerning the appropriate interpretation of the linguistic environments listed in Q8.

In sum, the results obtained on Q8 of the questionnaire must be interpreted with care. A more detailed description of each of the categories could possibly have led to different results. In addition, a different answer format rather than reporting numbers of weeks may have been appropriate for this question as well. Even if the frequencies of use that were reported for the ten language-based activities also depended on the participants' perceived amount of language use, the answers provided in the activity grid were expected to reflect their actual language use rather accurately, given that this perception was easier to rate according to more intuitive categories such as, e.g., 'several times a week', than by adding up numbers of hours or weeks dedicated to a given pastime. In spite of these critical considerations concerning its reliability, 'Exposure to English' was considered a complement to 'Frequency of use' and submitted to further analyses.

5.3.3   Correlations between L2 proficiency and exposure measures

Having thus defined the respective steps of variable formation for the two complementary exposure measures, correlational analyses were carried out to determine whether there was a relationship between proficiency in and exposure to English. Whereas English language proficiency, as measured by the MTELP, was not correlated with the secondary exposure measure, i.e., 'Exposure to English', a moderate and significant correlation was found with 'Frequency of use' ($r = .41$, $p = .001$), as illustrated in the Tables 22 and 23 below.

Table 22. *Correlations (Exposure to English / English proficiency)*

		Exposure to English	English proficiency
Exposure to English	Pearson Correlation	1	-.05
	Sig. (2-tailed)		.69
	N	66	66
English proficiency	Pearson Correlation	-.05	1
	Sig. (2-tailed)	.69	
	N	66	66

Table 23. *Correlations (Frequency of use / English proficiency)*

		Frequency of use: English	English proficiency
Frequency of use: English	Pearson Correlation	1	.41**
	Sig. (2-tailed)		.001
	N	66	66
English proficiency	Pearson Correlation	.41**	1
	Sig. (2-tailed)	.001	
	N	66	66

Note. ** $p > 0.01$.

Subsequently, the three subcategories of use derived from the principal component analysis discussed above, were also submitted to correlational analyses. Interestingly, only the frequencies pertaining to written interaction and reading, labeled 'Written use' ($r = .53$, $p < .001$), and those of interaction with a machine, labeled 'Technological use' ($r = .42$, $p < .001$) were significantly correlated with English proficiency. In contrast, there was a weak negative correlation between oral interaction and listening, labeled 'Oral use', and the MTELP scores ($r = -.12$, $p = .36$).

Naturally, the fact that the MTELP assessed receptive language skills that were measured exclusively in the written mode might explain the extent to which outside-of-school activities pertaining to written language would have a relatively high predictive value. At the same time, it is also reasonable to assume that the proficiency levels established by means of the MTELP cannot solely be explained on the basis of such personal language use. In fact, there are probably a number of classroom-based activities that were not listed in our questionnaire, that predict the kind of performance targeted in a language test, a considerable part of which pertained to knowledge of grammar. Still, it appears as if the dimension of 'Oral use' was inversely related to the proficiency measure. Even if this correlation was not significant, it seemed logical to infer that one possible reason why 'Exposure to English' was not related to English proficiency was the fact that this exposure measure, namely the overall amount of time spent in different English-speaking environments, probably referred to more oral than written interactions, on the whole. This

was at least partly confirmed by a moderate but significant correlation between 'Frequency of oral use' and 'Exposure to English' (Pearson's $r = .33$, $p < .01$). In other words, it may be that the oral component of the variable, which cannot be identified by merely specifying the linguistic environments (country, family, work, school) subsumed under 'Exposure to English', explains why there is no relationship between this exposure measure and the scores on the English proficiency test.

Finally, the analyses of the two exposure measures reveal their conceptual complementarity, on the one hand, and their methodological limitations, on the other. For purposes of triangulation, both variables will be entered into the regression model destined to determine possible predictors of positive lexical transfer from English. However, any results indicating possible relationships between either type of exposure and positive transfer will have to be interpreted with care, keeping in mind the above observations regarding reliability issues. Moreover, the subcomponents identified in 'Frequency of use' may provide a more detailed account of the role that L2 exposure and use might play in the activation of that language when manipulating a third.

5.4 Interest in the German language and culture

Finally, the last two variables that were investigated as possible predictors of positive transfer were the respective levels of interest in the target language and culture. As previously stated, these levels were determined on the basis of a five-point scale ranging from 'très faible' to 'très fort', thus indicating very low to very high levels of interest in the German language and culture respectively, precisely in Q19 and Q20 of the background questionnaire. In both categories, there were no ratings at the lowest level 'très faible' and only one participant rated their interest in the German language as 'faible'. Rather, the majority of the participants rated their interest in both German language and culture as strong ('fort') or very strong ('très fort'), as illustrated in the two pie charts below.

Figure 21. *Levels of interest in the German language*
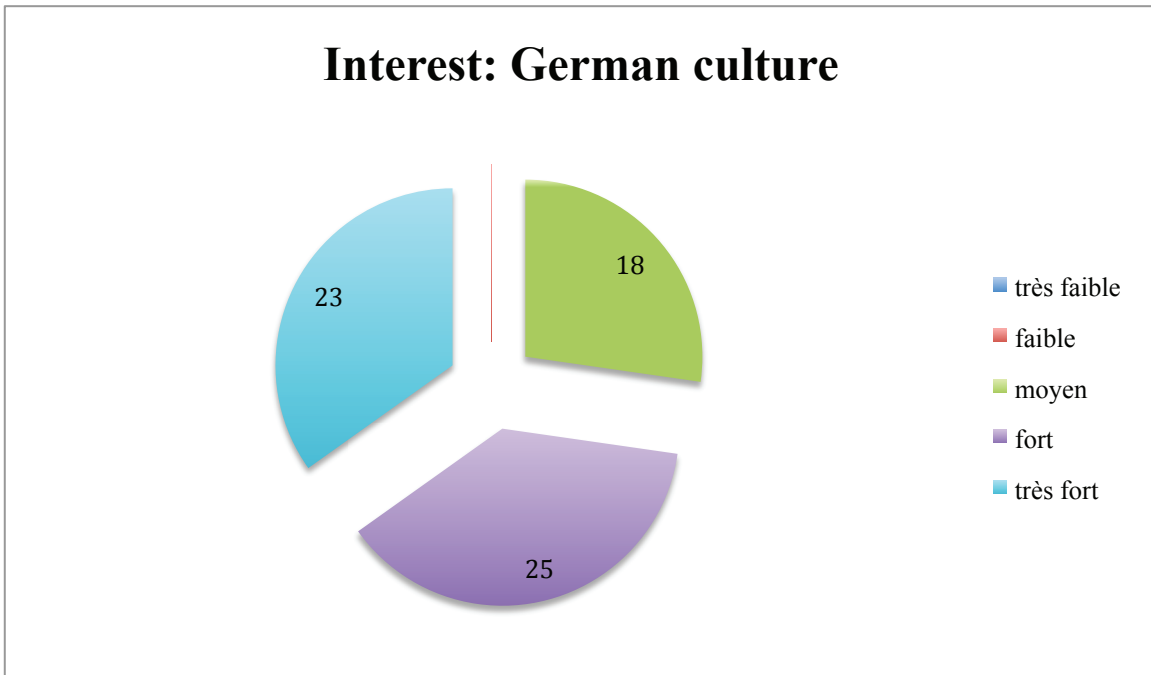
Figure 22. *Levels of interest in the German culture*

Given that all of the participants were enrolled in the same arts program with a language profile, this tendency towards higher levels of interest in the new target language and the related culture was not surprising. However, given that the specific program targeted two

foreign languages, namely German and Spanish, it was possible that several participants had made their choice to specialize in Spanish, yet without being particularly interested in learning German or vice versa. For purposes of comparison, the following subsection briefly elaborates on the relative levels of interest that were listed for either aspect of Spanish and English, respectively.

5.4.1 Levels of interest across languages

The two figures below were included to illustrate the participant ratings for Spanish in Q13 and Q14.
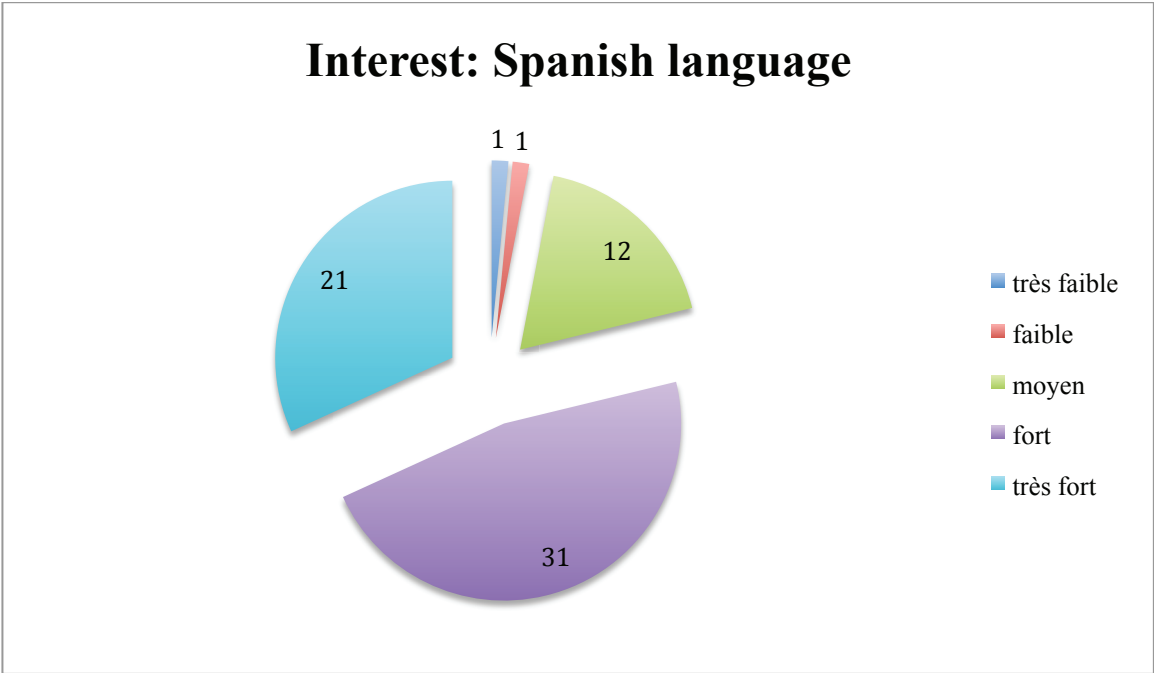
Figure 23. *Levels of interest in the Spanish language*

Interest: Hispanic culture

Legend:
- très faible
- faible
- moyen
- fort
- très fort

Values shown in chart: 1, 1, 18, 18, 28
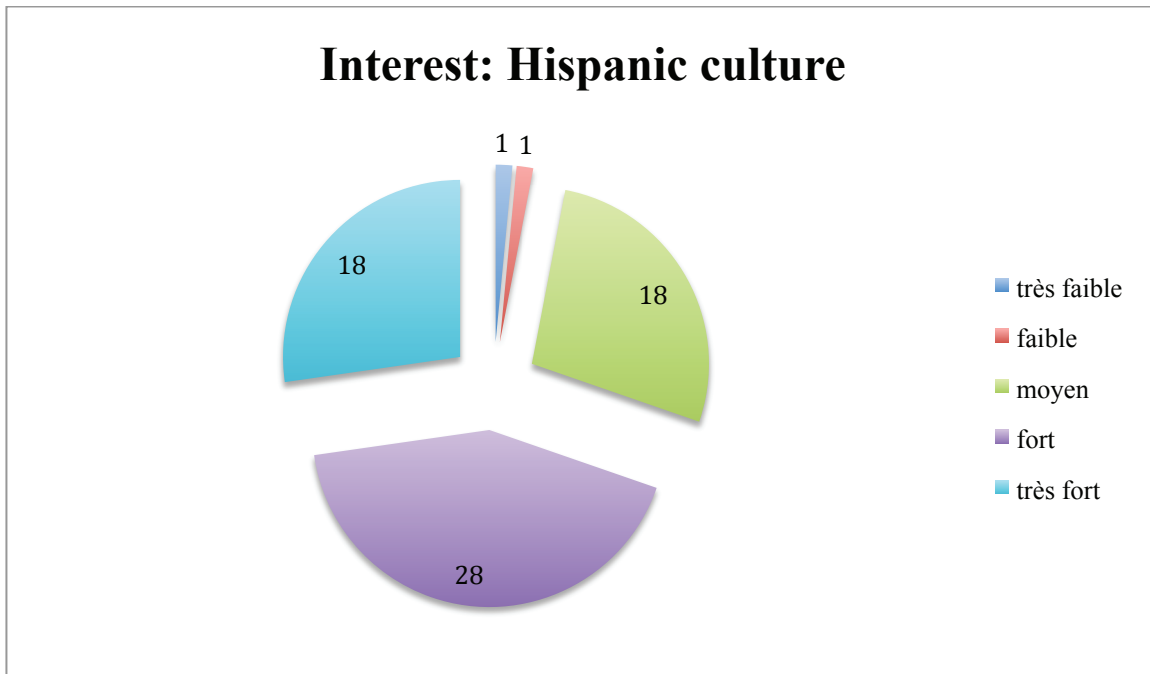
Figure 24. *Levels of interest in the Hispanic culture*

From this comparison, it follows that very high levels of interest in either language or culture were slightly more frequent for German than for Spanish. In addition, the distribution was largely comparable, the most frequent rating being 'fort', both across languages and categories. Since the same question had been posed for the second language English, we also included the respective ratings collected from Q9 and Q10. Even if the protocol had not initially been designed to investigate whether the participants were inclined to learn English or as to how far their degree of interest in their L2 might differ regarding the more recent target languages, Figure 25 illustrates that it was the English language that interested the largest number of participants.
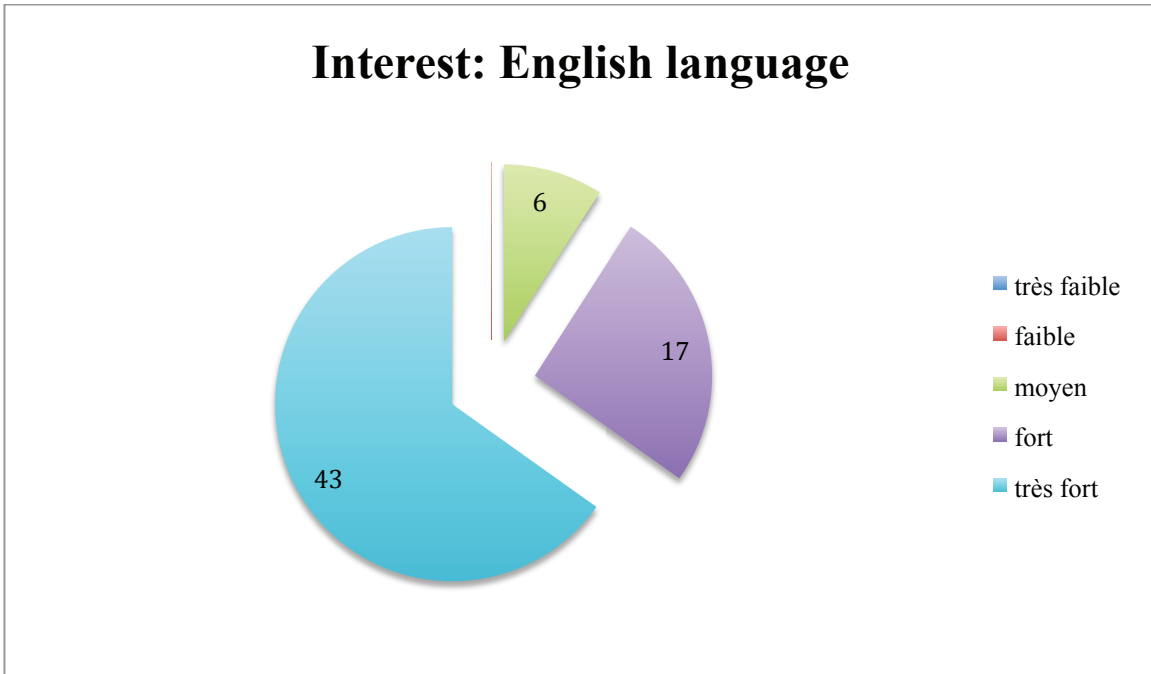
Figure 25. *Levels of interest in the English language*
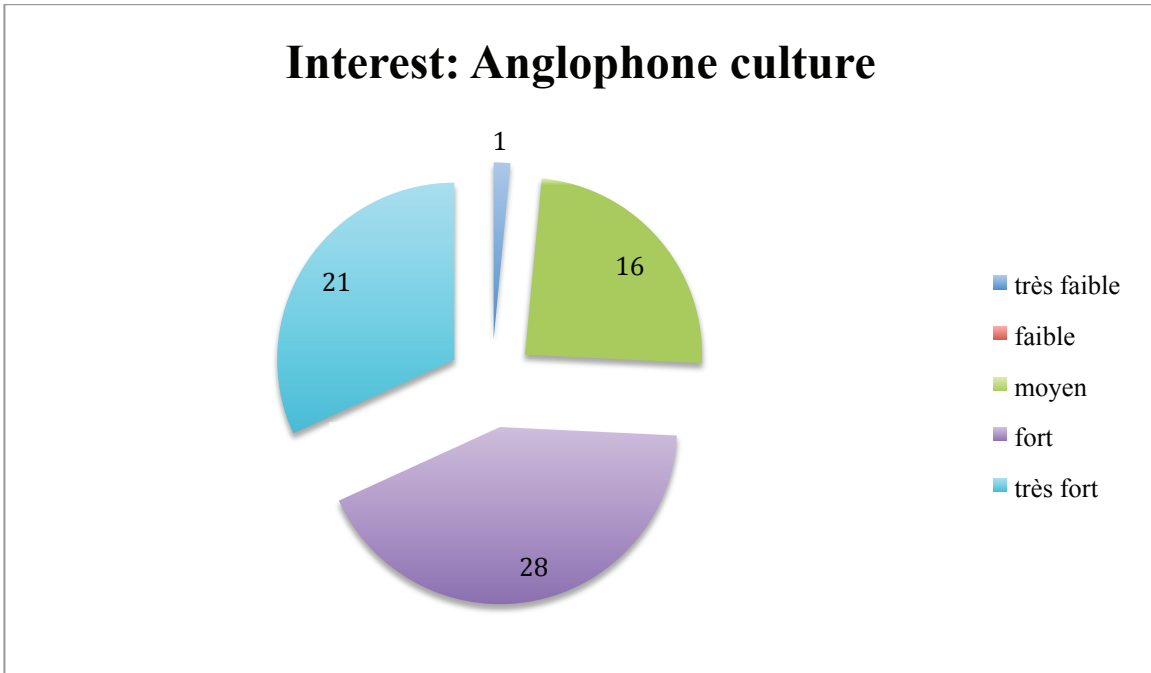
Figure 26. *Levels of interest in the Anglophone culture*

A total of 43 ratings were listed at the highest level, 'très fort'. Interestingly, this tendency for a strong interest in the English language was not consistent with the respective ratings regarding Anglophone culture (see Figure 26), where the distribution was again comparable

to the levels of interest listed for German and Hispanic culture. Even if the ratings were possibly biased by the fact that students enrolled in a language program might think they are supposed to have at least some sort of interest in their non-native languages, it was still remarkable that English was the language in which our participants appeared to be most interested. Of course, the assessment of this variable did not allow inferences as to the aspects of the language that motivated such favorable ratings. Most interestingly, these results are not consistent with the long-standing stereotypical representation of French-speaking Canadians concerning their attitudes towards English and vice versa, two linguistic communities that have been referred to as "two solitudes" (MacLennan, 1945). Rather, the overall positive perception of the English language in our sample appears to mirror more recent investigations into the attitudes of young francophone Quebeckers towards their L2. In fact, even if English is still seen as a threat to the French language by some French-speaking youth in the province of Quebec, many seem to recognize the value of learning English, be it for economic reasons or simply for the sake of its international prevalence (e.g., Remysen, 2004).

In sum, what should be retained from the above observations is that the results pertaining to levels of interest in the different non-native languages did not reveal much variability across the sample. In other words, relatively high levels of interest were listed for nearly all participants in all the languages and related cultures. However, this might be due to a lack of specificity of the measure, an issue that will be addressed in the following chapter, as a part of a critical look at the answer to this last research question (see section 6.1.5).

5.4.2 A restricted view of the affective dimension

As previously stated, the research protocol was constructed to account for different influential factors that were likely to impact positive lexical transfer from English to German in the present study. In accordance with Hufeisen's Factor Model (see Chapter 2), it seemed indispensable to include an affective dimension in our approach to multiple language learning. More precisely, the participants' inclination towards the target language itself and the communities that are typically associated with it was expected to influence their inclination towards learning that language. By extension, relative levels of interest in

the German language and culture, respectively, was likely to have an impact on the degree to which our participants would make an effort to understand the new words and thus to be more or less inclined to mobilize various resources to make sense of the target items. Our last research question was motivated by this reflection (see section 3.5).

As suggested by the results presented above, the majority of our participants have strong or even very strong interest in the German language and culture. In contrast, the outcome of the translation task regarding both transfer rates and levels of crosslinguistic awareness was highly variable. Prior to analyzing the statistical contribution of this variable within the regression model presented in the following section, it seems important to point to the way of operationalizing 'interest', as opposed to the other predictive variables that have been presented and discussed above. In fact, even if the concept of interest, such as it was presented in the background questionnaire, was not expected to be misinterpreted, the five-point scale ranging from 'very low' (1) to 'very high' (5) might have been insufficient to make out differences across the sample. Since the participants were enrolled in the same language program, it may have been preferable to proceed to a finer-grained analysis of this affective dimension. Unfortunately however, it was beyond the scope of the present study to increase the number and volume of materials administered to our participants. It must be acknowledged that the questions pertaining to levels of interest in different languages and related cultures could not generate answers that would provide a detailed account of our participants' attitudes towards the target language communities and their language. In turn, the present research protocol could not provide any substantial observations regarding the affective dimension of L3 learning.

5.5   Predicting positive lexical transfer from English

As briefly outlined in Chapter 4, we ran binomial logistic regressions to investigate possible predictors of positive transfer. The choice of this statistical model was motivated by the fact that the response variable was assessed by means of the number of successful trials. That is, each of the German test items that could be rated as transfer or non-transfer from English represented one of 28 trials. However, given that some participants had prior knowledge of certain test items, the response variable was normalized, i.e., total transfer

rates for each participant were calculated in relation to the number of items that were in fact unknown. In other words, the total number of trials varied from one participant to another, after instances of prior knowledge were excluded.

5.5.1 Binomial logistic regressions

In order to answer research questions 2 to 5, the following influential factors were examined regarding their relationship to positive transfer from English:

- Metalinguistic awareness (MLA)
- English proficiency
- Exposure to English
- Interest in the German language and culture

As indicated in the respective sections of the present chapter, two of these concepts were assessed by means of two complementary measures. Namely, the primary measure of MLA was the THAM, which was complemented by the metalinguistic ratings of the TAPs (see 5.2), while the primary exposure measure was 'Frequency of use', which was complemented by 'Exposure to English' (see 5.3). Accordingly, the following variables were entered into the equation as covariates, in terms of potential predictors of 'Transfer from English', the dependent variable:

- 'MLA_TAP_Total' = MLA score from TAPs
- 'THAM_Total' = MLA score from THAM
- 'MTELP_Total' = English proficiency
- 'UseEng_Total' = Frequency of use (English)
- 'ExposureEng_Total' = Exposure to English
- 'Int_GermLang' = Interest: German language
- 'Int_GermCult' = Interest: German culture

Table 24 simply serves as a reminder of the descriptive results of each of these measures, as exposed in the previous subsections of this chapter.

194

Table 24. *Continuous Variable Information (Covariates)*

		Min	Max	Mean	SD
Covariate	MLA score from TAPs (raw) (112)	17	74	44.41	12.68
	MLA score from THAM (48)	6	37	18.36	7.09
	English proficiency (100)	28	99	68.09	16.81
	Frequency of use: English (40)	7	40	21.44	7.40
	Exposure to English (16)	0	12	2.38	2.78
	Interest: German language (5)	2	5	4.23	.74
	Interest: German culture (5)	3	5	4.08	.79

In addition, variance inflation factors (VIF) between 1.21 and 1.72 indicated that none of the explanatory variables were strongly correlated between each other. In other words, there was no multicollinearity among the independent variables.

5.5.1.1 Regression model featuring the THAM as the only measure of MLA

As briefly discussed with respect to the coding procedures of the metalinguistic dimension of the TAPs (see sections 4.5.1.1. and 5.2.2.2), there are some conceptual limitations regarding the inclusion of the MLA score from the TAPs as a predictive variable of positive transfer. In fact, the two categories labeled 'Absence' and 'Prior knowledge' according to the metalinguistic dimension logically implied non-transfer. In turn, the fact that the respective ratings were zero (0) in both dimensions engendered a sort of interdependence between the measures of 'Transfer from English' and the MLA score from the TAPs. As a consequence, the primary aim being to create a robust model to expose conceivable relationships between potential predictors and the response variable, the first logical step was to run the regression analysis without entering the secondary measure of MLA.[58] Table 25 illustrates the parameter estimates of this first regression model, in which 'THAM_Total' was the exclusive measure of MLA.

---

[58] For this particular statistical test, we used the SAS software instead of SPSS, given that a certain number of elements pertaining to the predictive power of the model could thus be directly extracted, as will be illustrated and discussed in the following subsection.

Table 25. *Parameter estimates (First model)*

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
			Analysis of Maximum Likelihood Estimates		
Intercept	1	-1.05	.37	7.91	.005
THAM_Total	1	.02	.01	8.83**	.003
MTELP_Total	1	.01	.00	6.62*	.01
UseEng_Total	1	.00	.01	.27	.60
ExposureEng_Total	1	-.03	.02	1.57	.21
Int_GermLang	1	.10	.08	1.37	.24
Int_GermCult	1	-.16	.08	3.69	.06

Note. * $p < .05$. ** $p < 0.01$.

The primary information to be retained from these statistical results is the column labeled Wald Chi-Square. In particular, the highest statistic was reported for THAM_Total ($X^2 = 8.83$) thus indicating that MLA was the strongest predictor of positive transfer, followed by English proficiency ($X^2 = 6.62$). The last column refers to the probability of obtaining the Chi-Square statistic if there was no effect of the independent variables, i.e., if the null hypothesis were true. As illustrated by the respective *p*-values listed for each of the independent variables of this model, both THAM_Total and MTELP_Total could predict positive transfer significantly. In turn, none of the remaining variables appeared to have a positive impact on the response variable. Most surprisingly however, the estimate of both ExposureEng_Total and Int_GermCult, was negative, thus indicating the possibility of an inverse relationship with the response variable. In the case of the former, it appears as if this relationship was negligible ($p = .21$), whereas the latter nearly reached the level of significance ($p = .06$). Possible explanations for this finding will be presented and discussed in the Chapter 6. Finally, the lowest statistic was listed for UseEng_Total ($X^2 = .27$), thus suggesting that the frequency of English language use outside the classroom was not in the least related to the rate at which participants resorted to their L2 to translate new lexical items in German. This observation will also be interpreted in the following chapter when elaborating on the results with respect to each of the research questions, individually.

5.5.1.2 Regression model including the MLA score from the TAPs

In order to substantiate the above findings regarding the role of MLA in positive lexical transfer, the metalinguistic processing data extracted from the TAPs were submitted to subsequent analyses, with the aim of including this secondary measure of MLA in the equation. At first, we determined the proportion of instances where MLA and transfer measures were interrelated, i.e., where the total number of instances rated at the level of 'Absence' (N = 321) and 'Prior knowledge' (N = 120) represented less than 24 % of the totality of the ratings (N = 1848). Given that the remaining data pertaining to the levels of I, N, MA and UR were considered independent of the transfer ratings, we computed a variable that was based exclusively on the MLA ratings that were different from (0). To do so, for each participant we divided the total MLA score from the TAPs by the number of instances rated non-zero. In other words, we calculated the mean of the non-zero ratings. Consequently, the variable thus labeled 'MLA_TAP_Mean_nonzero' could be considered independent from 'Transfer_Total'. Hence, we proceeded to a subsequent analysis in which the MLA score from the TAPs was entered into the equation. As illustrated in Table 26, this step appeared to change the overall picture substantially. Finally, this normalized MLA score from the TAPs turned out to be the strongest predictor of positive transfer (X^2 = 26.44, p = .000), followed by the MLA score from the THAM (X^2 = 4.55, p = .03). Another crucial observation concerning the statistical results of this second regression model is that by adding the MLA score from the TAPs, English language proficiency does not qualify as predictor anymore (X^2 = 3.12, p = .08).

Table 26. *Parameter estimates (Second model)*

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
				Analysis of Maximum Likelihood Estimates	
Intercept	1	-1.95	.41	22.23	<.0001
MLA_TAP_Mean_nonzero	1	.76	.15	26.44**	<.0001
THAM_Total	1	.02	.01	4.55*	.03
MTELP_Total	1	.01	.00	3.12	.08
UseEng_Total	1	.00	.01	.08	.78
ExposureEng_Total	1	-.02	.02	.72	.40
Int_GermLang	1	.03	.08	.11	.74
Int_GermCult	1	-.17	.08	4.18*	.04

Note. $* p < .05. ** p < 0.01.$

A critical look at the statistics of the remaining variables suggests that the negative impact of Int_GermCult was slightly amplified in this model ($X^2 = 4.18$, $p = .04$), whereas UseEng_Total remained at the bottom of the list, as far as possible relationships to the response variable were concerned. Prior to discussing these findings in the last subsection of this chapter, it is important to examine the predictive power of this second regression model, in comparison to the previous analysis, in which the THAM was the exclusive measure of MLA.

5.5.1.3 Predictive power of the models

In linear multiple regression, researchers commonly refer to R^2 values to determine how well the model can predict the dependent variable based on the independent variables. In other words, this statistic helps determine what part of the variance the model can account for. Logistic regression does not have an equivalent to this R^2 value. Even if there are different kinds of pseudo-R^2 statistics, they should be interpreted with caution, given that low R^2 values, which appear to the norm in logistic regression, can induce misinterpretations regarding the proportion of variance explained by the predictors (see Hosmer & Lemeshow, 2000, p. 167). As a consequence, it was judged more reasonable to resort to another type of statistical analysis, precisely the area under the ROC-curve, in

order to determine the predictive power of each of the two regression models. Tables 27 and 28 report different association statistics based on the relationship between pairs of observations for each of the respective regression models.

Table 27. *Association Statistics (First model)*

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	57.1	**Somers' D**	0.16
Percent Discordant	40.8	**Gamma**	0.17
Percent Tied	2.0	**Tau-a**	0.08
Pairs	739935	**c**	0.58

Table 28. *Association Statistics (Second model)*

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	59.9	**Somers' D**	0.22
Percent Discordant	37.9	**Gamma**	0.23
Percent Tied	2.3	**Tau-a**	0.11
Pairs	739935	**c**	0.61

The concordance index *c* in each of the tables above indicates that the area under the ROC curve was either slightly below (Table 27) or slightly above (Table 28) 60 %, which represents the percentage of randomly drawn pairs for which the predictions made by either model were true. Even if indices between .6 and .7 are still considered "poor" with respect to the accuracy of diagnostic tests in medical science (Tape, n.d.), investigations into human behavior can naturally be expected to be less accurate. Finally, with reference to the slightly lower concordance index that was listed for the first regression model, in which the THAM was the exclusive measure of MLA (*c* = .58), it can be assumed that the complementary measure of MLA could account for a part of the variance. In sum, the regression analysis that included both measures of MLA appeared to predict positive transfer more accurately. Furthermore, a Hosmer and Lemeshow goodness-of-fit test was performed to examine whether each of the models was also consistent with the data. In this type of statistical test, binomial data are grouped into so-called "profiles" that have the

same values on the predictor variables. Subsequently, the observed number of events and non-events is compared to the expected number by means of a Chi-Square analysis. The two tables below (29 and 30) illustrate the Hosmer and Lemeshow test results for each of the models respectively.

Table 29. *Hosmer and Lemeshow Goodness-of-fit test (First model)*

Chi-Square	DF	Pr > ChiSq
9.49	8	0.30

Table 30. *Hosmer and Lemeshow Goodness-of-fit test (Second model)*

Chi-Square	DF	Pr > ChiSq
14.72	8	0.07

In this kind of analysis, low *p*-values are indicative of poor fit: "Based on the X^2 distribution, a Hosmer-Lemeshow (H-L) statistic with a *p*-value greater than 0.05 is considered a good fit. The lower the H-L statistic, the less variance in fit, and the higher the *p*-value" (Hardin & Hilbe, 2007, p. 131). Finally, even if the second model revealed more variance, as indicated by the higher Chi-Square statistic ($X^2 = 14.72$), both regression models appeared to be correctly specified, given that the observed threshold for each of them was above .05 (see Tables 29 and 30).

5.5.2 Positive transfer in relation to each of the predictor variables

In order to provide a visual illustration of the relationship between positive transfer and each of the predictor variables identified in the regression models, this section provides an overview of the correlations between positive transfer and each of its potential predictors. Prior to discussing the significant relationships in more detail, Table 31 illustrates the complete correlation matrix for the whole data set.

Table 31. *Correlations between variables*

	1a	1b	2	3	4	5	6	7
Pearson correlation matrix								
1) Translation task								
a) Transfer from English	--	$.563^{**}$	$.464^{**}$	$.403^{**}$	.082	-.127	.036	-.142
b) MLA - TAP		--	$.264^{*}$	$.267^{*}$	.125	-.099	$.256^{*}$	.175
2) THAM-3			--	$.360^{**}$	.003	-.087	.157	.006
3) L2 proficiency (MTELP)				--	$.408^{**}$	-.032	.086	.020
4) Frequency of L2 use					--	$.272^{*}$	.105	.136
5) L2 Exposure						--	-.028	-.237
6) Interest: German language							--	$.575^{**}$
7) Interest: German culture								--

Note. $* p < .05. ** p < 0.01.$

Apart from providing visual support of the significant correlations between the response variable and its main predictors, Table 31 also indicates how the variables are related to each other. Moreover, given that two measures of this data set had been established on the basis of subscales, it seemed relevant to insert an additional correlation matrix. Thus, to provide a more detailed account of the findings, Table 32 represents the complete correlation matrix, in which the relative contribution of the L and ML scores of the THAM-3, as well as the respective subsections of the MTELP, i.e., grammar, vocabulary and reading comprehension, are listed separately.

Table 32. *Correlations between variables, including subscales of THAM-3 and MTELP*

Pearson correlation matrix

		1a	1b	THAM-3		MTELP			4	5	6	7
				L	ML	Gram	Voc	Read				
1)	**Translation task**											
a)	**Transfer from English**	--										
b)	**MLA - TAP**	.563**	--									
2)	**THAM-3**											
a)	**L score**	.331**	.186	--								
b)	**ML score**	.470**	.270*	.583**	--							
3)	**L2 proficiency (MTELP)**											
a)	**Grammar**	.386**	.273*	.276*	.307*	--						
b)	**Vocabulary**	.426**	.304*	.344**	.259*	.813**	--					
c)	**Reading**	.243*	.142	.303*	.319**	.716**	.760**	--				
4)	**Frequency of L2 use**	.082	.125	.138	.244*	.428**	.363**	.383**	--			
5)	**L2 Exposure**	-.127	-.099	.055	-.088	.006	-.091	.033	.272*	--		
6)	**Interest: German language**	.036	.256*	.217	-.165	.070	.125	.022	.105	-.028	--	
7)	**Interest: German culture**	-.142	.175	.098	.089	-.044	.079	.043	.136	-.237	.575**	--

Having thus provided an overview of the intercorrelations for the whole data set, the following three subsections are devoted to the specific relationships between positive transfer and its predictor variables, including scatterplots of the relative distribution of the test results. Note that the transfer measure was normalized in all analyses, as previously explained (see sections 5.1.1 and 5.5). Similarly to the normalized MLA score from the TAPs, which excluded all zero-ratings, the normalized transfer score excluded ratings that indicated prior knowledge. As will be discussed with regard to each of the scatterplots, this normalization entailed inflations of the total score for participants with high transfer rates in combination with prior knowledge of a certain number of test items.

5.5.2.1 Relationship between the reflexive dimension of MLA and positive transfer

The scatterplot in Figure 27 illustrates the relative distribution of the THAM scores in relation to the normalized transfer rates. As previously stated, all the correct translations based on prior knowledge were extracted from the transfer measure and the final scores normalized across the sample.
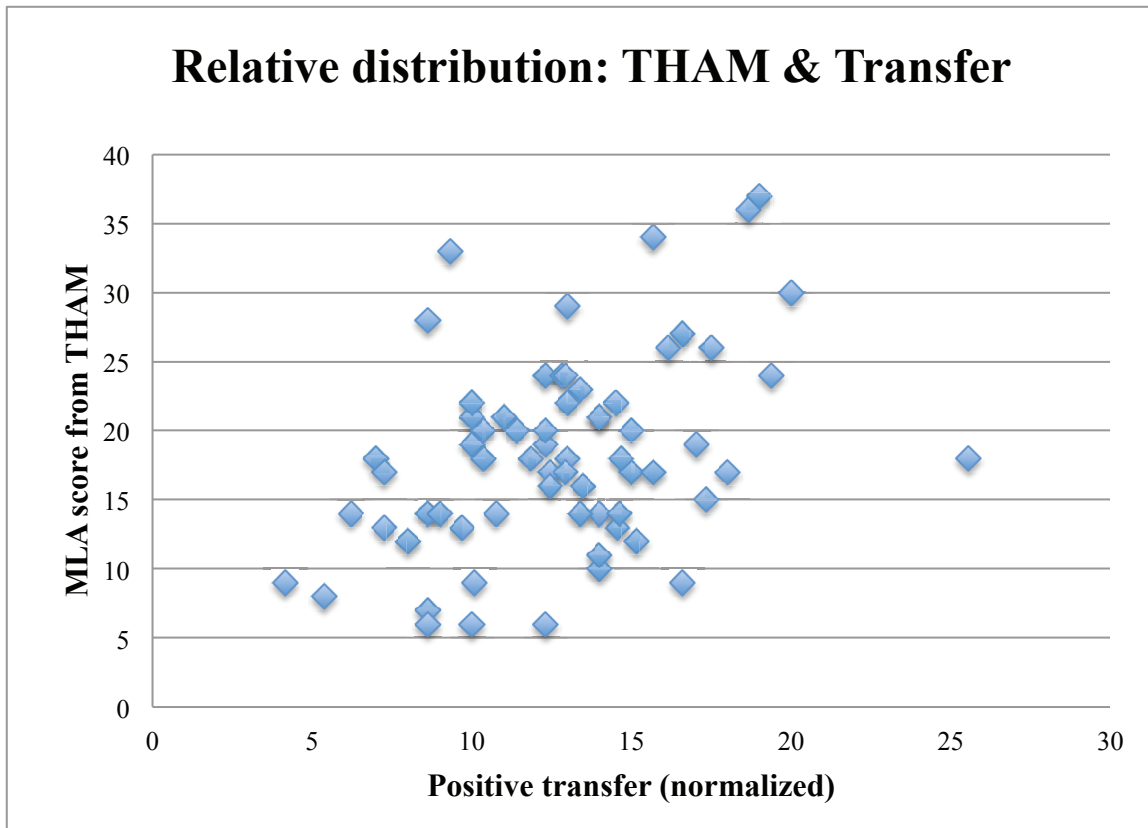
Figure 27. *Relative distribution (MLA score from THAM / Positive transfer)*
Note. The normalized transfer measure excludes *Prior knowledge* (PK) ratings.

The scatterplot above (Figure 27) reveals a positive slope, thus corroborating the relationship between the two variables, identified by means of the regression model discussed above. Not surprisingly, the correlation between the THAM score and the normalized transfer measure was highly significant (r = .46, p < .001). However, there appears to be an outlier in the mid-right area of the graph, who was not discernable when entering raw instead of normalized scores. What must be taken into consideration is that that prior knowledge of a number of the target items somehow advantaged learners on their normalized transfer score. This is particularly evident for one participant in the plot above (see Figure 27), who seems to stand out from the rest of the group. In fact, her transfer rate (ID #58) was inflated due to the fact that 5 out of the 28 German-English cognates were already known. Even though, technically speaking, normalization should not affect the general tendency, it must be acknowledged that some associations to be established between target items and their English cognates were less obvious than others, due to different degrees of formal similarity. Thus, it makes a qualitative difference which words

were already known, given that the chances to transfer the remaining unknown items tend to be higher for participants with prior knowledge of difficult words. In fact, it may be worth considering the relative effort of the participant with the highest transfer rate (ID #58, TR = 21) in relation to the specific words that she already knew, among which were *Brust* and *Buch*, for example.[59] As suggested by the respective frequencies with which these items were transferred across the sample (*Brust* = 1, *Buch* = 25, see Figure 2), prior knowledge of these particular words could clearly have been an advantage, given that the chances to get them right if they had been unknown, seemed relatively low. What should be retained from these considerations is that prior knowledge could not simply be ruled out by normalization, given that it did not only affect the results quantitatively but also qualitatively. In other words, these normalized scores must be interpreted with care, since quantitative results alone cannot account for the variability within and among participants, which further points to the usefulness of data triangulation. On the whole, the above observations are consistent with our hypotheses in so far as our results indicate a relationship between positive lexical transfer from English and the reflexive dimension of MLA, as evidenced by the results of the THAM.

5.5.2.2 Relationship between the applied dimension of MLA and positive transfer

The relative distribution of the normalized MLA score from the TAPs in relation to the normalized transfer rates is illustrated in Figure 28. As discussed in the previous chapter, we extracted the instances identified as *Absence* or *Prior knowledge*, both of which were rated 0, which created an overlap between the MLA and the transfer measure.

---

[59] Note that ID #58 was not one of the eight participants who were added to our sample subsequently.
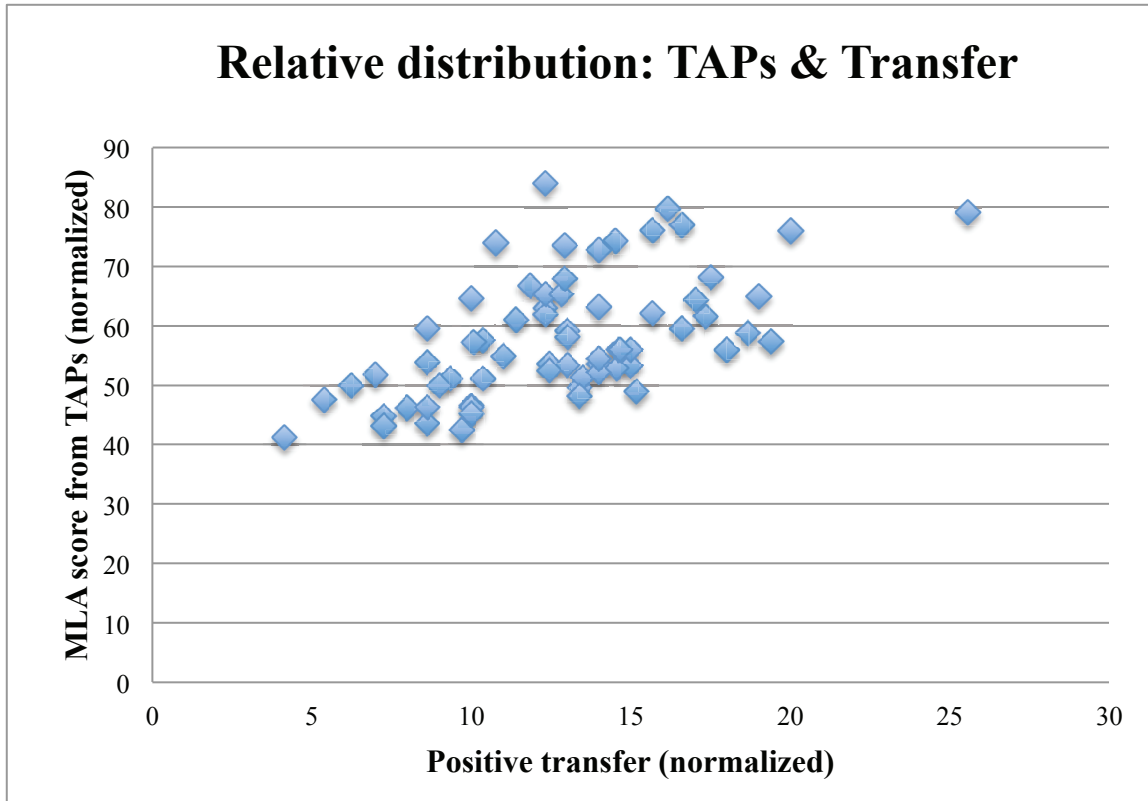
Figure 28. *Relative distribution (MLA score from TAPs / Positive transfer)*
Note. The MLA score from the TAPs was normalized for both zero-ratings *Absence* (A) and *Prior knowledge* (PK). The transfer score was normalized for *Prior knowledge* (PK).

From a comparison of the distribution illustrated in Figure 28 to the distribution of the raw scores featured in Figure 15 (see section 5.2.2.2), two observations arise. First, the slope remains positive, thus indicating a relationship between the two variables. This was confirmed by a significant correlation ($r = .56$, $p < .001$), which was even slightly stronger than the one between the raw scores ($r = .53$, $p < .001$). Second, we must keep in mind the point that was raised about the normalized transfer measure in the previous section. In fact, the same participant whose particularly high transfer rate (TR = 21) was inflated with the prior knowledge index (see previous section) also appears like an outlier in the top right corner of this graph (Figure 28). Of course, it would have been desirable that none of the target items were known to any of the participants at the time of data collection. However, as previously stated, it was impossible to control for this factor altogether. Even if the participants had not received any instruction in German, they may already have come across one or the other of the target items in another context. Of course, the above

observation must be considered a weakness of the translation task, and this aspect will be further discussed among the limitations of the study (see section 6.3). Still, the results clearly indicate that higher levels of crosslinguistic awareness, such as evidenced by the MLA from the TAPs, were strongly correlated with the transfer measure.

5.5.2.3 Relationship between English proficiency and positive transfer

The above observations underpin our initial assumptions concerning the crucial importance of resorting to metalinguistic analyses of target items and structures in order to come to the right conclusions about the *potential vocabulary* that was activated from a related background language, either implicitly as a sort of intuition or explicitly by an active word search across languages. Even though proficiency in the source language from which these interlingual identifications are drawn, does affect transfer rates, higher levels of English language proficiency did not necessarily lead participants to resort to positive transfer, as illustrated in the top left corner of the scatterplot below (Figure 29).
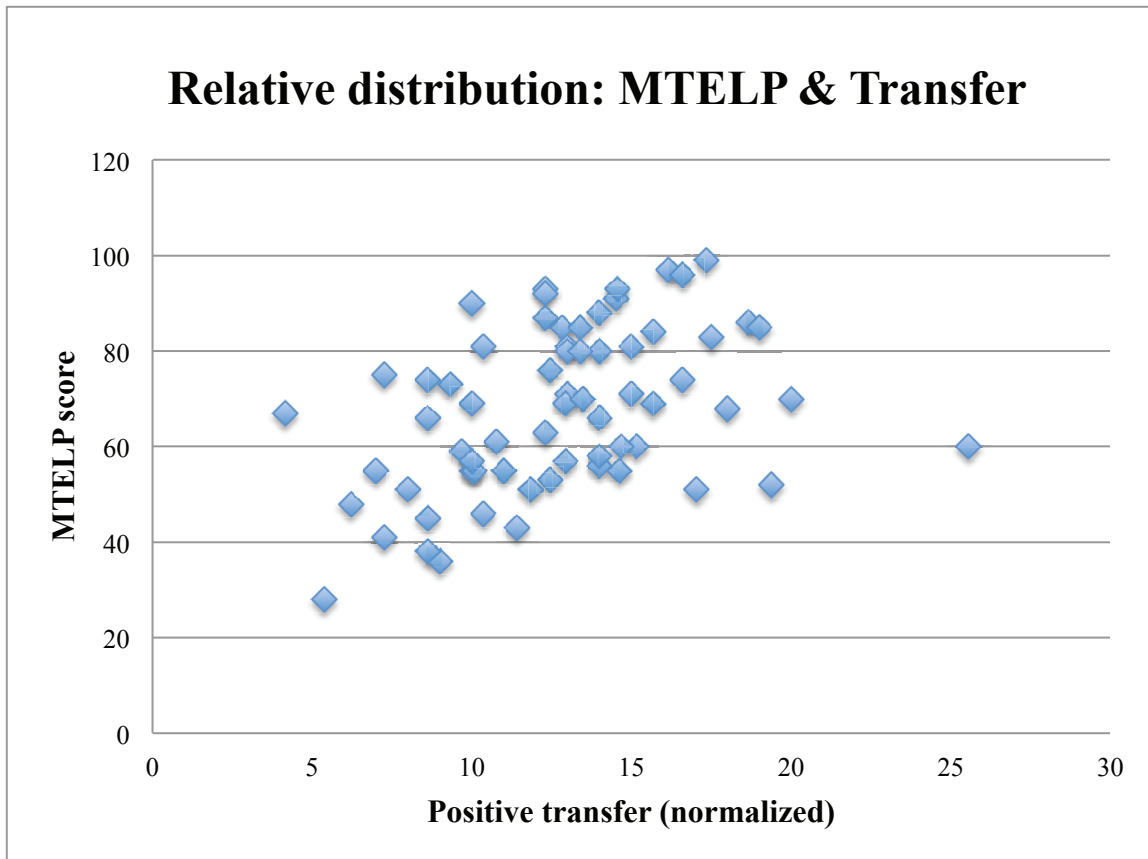
Figure 29. *Relative distribution (English proficiency / Positive transfer)*

However, it must be stated that the levels of L2 proficiency were relatively high throughout the sample (Min = 28, Max = 99, Mean = 68.9, Median = 69) and that all of the participants appeared to have sufficient knowledge of English to establish links between the new words in German (L3) and the corresponding background vocabulary in English (L2). Thus, as suggested with regard to lexical access in speakers of multiple languages, there might be a threshold level of proficiency for a language to compete in the selection process (de Bot, 2004, p. 8). Considering that our participants had received between 6 and 13 years of English instruction[60], that the vast majority of them attained scores above 60 on the MTELP and that all of them used their English at least in some of the listed non-instructional settings and very regularly at school, it can be assumed that the language as a whole was at a relatively high level of activation. However, the fact that the L2 was present

---

[60] The majority of the 66 participants had started to learn English in third year of primary school around the age of 8 or 9. However, two of the 17-year olds reported that they were only 11 years old when they started to learn English. In contrast, two of the older participants had started in first year of primary school, which explains the large gap. All the participants had been schooled in the province of Quebec.

in the learners' minds and that they were capable of manipulating intermediate to advanced grammar, vocabulary and reading comprehension issues in their L2 did not appear to be the primary condition for them to make active use of their L2 background vocabulary in an L3 learning context. That is, high scores on the MTELP did not always relate to high transfer rates (see Figure 29). However, the positive slope in both figures indicates a relationship between the MTELP and the transfer measure. Pearson correlations between the two variables were significant ($r = .40$, $p < .001$).

5.5.3 Overall results

This last section of the chapter focused on the statistical analyses that were applied to identify the variables that qualified as predictors of positive lexical transfer from English. Two models of binomial logistic regression were examined to investigate these potential relationships. In the first model, the only measure of MLA was the THAM, whereas the second model also included the normalized MLA score based on the analysis of the TAPs. Even if this complementary measure of MLA was manipulated in order to eliminate the overlap with the measure of transfer, it still must be born in mind that both variables were derived from the same data, i.e., the verbalizations produced during the translation task. As a consequence, the results from the second regression model must be interpreted with caution. Most importantly however, the observed test results from both regression analyses qualified MLA as the strongest predictor of positive transfer from English. When the THAM was the exclusive measure of MLA, English proficiency was the second variable that could significantly predict positive transfer, yet with a slightly lower statistic, i.e., $X^2 = 6.62$, as opposed to $X^2 = 8.83$ for the THAM. When the normalized MLA score from the TAPs was entered into the equation, the latter became the strongest predictor ($X^2 = 26.44$, $p < .001$), followed by the MLA score from the THAM ($X^2 = 4.55$, $p = .03$), whereas English language proficiency no longer qualified as a significant predictor ($X^2 = 3.12$, $p = .08$).

On the whole, these results suggest that a finer-grained methodological approach to the concept of MLA might lead to an amplification of the crucial role it has been assigned in multiple language learning. Namely, the two complementary measures of MLA that have been used in this study to operationalize two different aspects of MLA, i.e., a reflexive and

an applied dimension, appeared to be the most influential factors regarding the active use of English (L2) vocabulary knowledge to translate unknown words in German (L3). In turn, none of the variables pertaining to L2 exposure or frequency of use predicted such positive transfer. Moreover, whereas levels of interest in the German language did not appear to bear any relationship with the response variable, higher levels of interest in the German culture seemed to be associated with lower transfer rates. A critical analysis of these results will be provided in the following chapter, where each of the research questions will be answered, prior to a comprehensive discussion of the contribution of the present study to the presumed role of MLA in the L3 learning process.

CHAPTER 6

6 Discussion

The goal of the present research was to determine the variables that could significantly predict positive lexical transfer from English (L2) to German (L3). Besides the most discussed influential factors such as source language proficiency and exposure, the present study attempted to isolate MLA as a potential predictor of positive transfer and thus to firmly establish its role in this active process of crosslinguistic interaction between two non-native languages that are typologically related. In the first subsection of the present chapter, we will elaborate on the relationship between each of the predictive variables and positive transfer, namely by discussing the results with respect to our research questions. From this summary of our findings, we will proceed to a critical analysis of the role of MLA in the L3 learning process, such as it has been established in the present study, with reference to current research on multilingualism and third language acquisition.

6.1 Positive transfer and its potential predictors

In the particular context of our study, positive transfer was operationalized by the frequency with which French-speaking Quebeckers resorted to their L2 English to translate a series of mostly unknown lexical items in their L3 German. Having established the presence of positive lexical transfer as a translation strategy in the present study (see section 6.1.1), the influential factors that were investigated as potential predictors of such positive lexical transfer were our participants' levels of metalinguistic awareness (MLA), their levels of English language proficiency and exposure, as well as their levels of interest in the German language and culture. Below, our findings relative to each of these predictive variables will be discussed with reference to current research perspectives related to these issues.

6.1.1 Do French-speaking learners of German (L3) make use of their prior knowledge of English (L2) to discover the meaning of new lexical items in German (L3), thus revealing positive lexical transfer from English?

Our first research question addressed the notion of lexical transfer itself. The analysis of a

word translation task in combination with think-aloud protocols revealed that the 66 participants in our sample resorted significantly to positive transfer from English to translate mostly unknown words from German (L3) into French (L1). As outlined in the previous chapter, the total number of transferred items across the sample was of 783 out of 1848 translation attempts for the 28 test items that could possibly be transferred on the basis of an English translation equivalent. This number represented a mean of 11.86 transferred items out of 28. The 95% confidence interval ranging from 11.01 to 12.72 did not include zero, thus indicating that transfer does in fact occur. This fundamental observation served to answer our first research question, namely positive lexical transfer from English could be identified as a comprehension strategy among the participants of our sample. Moreover, the extent to which cognate vocabulary from this typologically related background language was activated to make sense of the target items appeared to be related to the degree of similarity between the given target and its English translation equivalent.

This observation is in accordance with a large body of research in second and third language acquisition (e.g., Ringbom & Jarvis, 2009), stating that transfer phenomena are most common where crosslinguistic correspondences are most salient. Moreover, our findings corroborate previous research in foreign language vocabulary acquisition, which holds that initial stage learners tend to rely on formal representations of words, which often results in item transfer, whereas procedural (or meaning-based) transfer (see section 2.4.1) appears to become more frequent with an increase in TL proficiency (e.g., Lindqvist, 2010). However, it must be acknowledged that procedural transfer has mostly been reported in production data where learners were shown to transfer semantic properties of words, such as in phrases like "*animaldoctor*" to refer to a *veterinarian*, based on the Finnish translation equivalent *eläinlä ä käri* that literally means 'animal doctor' (Meriläinen, 2006, as cited in Jarvis, 2009, p. 115). Likewise, transfer of syntactic features such as "She kissed *with* him" (Jarvis, 2009, p. 117), which is also considered procedural, generally refers to instances of deviant TL use. Hence, two basic characteristics of our protocol make a clear-cut categorization difficult. First, we focused on positive transfer, where the mental access routes for a given target structure are less easily detected than in instances of negative transfer. Second, we focused on the comprehension of pre-selected items and structures,

which partially predetermined the sort of associations that our participants were going to make. Moreover, given the specific task type that was chosen for the present study, the kind of lexical transfer that we observed was mostly item-based. More specifically, the fact that our participants were presented with a series of isolated words naturally triggered mainly form-based associations, i.e., similarities in spelling and sound.

However, typological similarities do not in themselves represent a sufficient condition for a learner to establish potentially helpful links between the corresponding structures in their different languages. Otherwise, all of the test items with high-similarity cognates should have been transferred throughout the sample, which was not the case. Another language-inherent feature that was examined regarding its impact on the transfer rates of individual test items was that of word frequency. Whereas it may have been expected that high frequency words in English would be activated more easily as a basis of lexical transfer, our results did not indicate such a tendency (see section 5.1.3.2). From a usage-based perspective, this could be explained in terms of construction learning. That is, learners do not encounter lexical items in an isolated fashion but mostly as parts of constructions. As a consequence, it is not the frequency of lexical items alone, but also that of the constructions in which they commonly occur, that influences perception and memory. Moreover, in a recent contribution on usage-based language learning, Nick Ellis and his colleagues remind us of longstanding research perspectives in psychology emphasizing the crucial importance of contingency between cue and interpretation for associative learning (N. C. Ellis, O'Donnell, & Römer, 2015, p. 169). In other words, besides the mere frequency of usage to which learners can potentially be exposed, associations also depend on the specific types of relationships that can be established with other learning experiences. It is not because a particular word has been encountered more often that it is necessarily processed in association with a given cue, in this case a similar item or structure from another language.

In short, our first hypothesis stating that the participants of our study would resort to their second language English to understand new words in their third language, German (see section 3.5) was confirmed. Moreover, our findings point to the fact that language-inherent factors such as typological proximity, as evidenced in the presence of cognate vocabulary,

or word frequency in either source or target language, cannot in themselves predict the extent to which learners resort to positive transfer. In line with a substantial body of research on transfer phenomena in multilingual contexts, we adopted a psycholinguistic approach to investigate different learner variables that were expected to affect the processing of unknown L3 (German) words by activating cognate vocabulary in a related L2 (English).

6.1.2 Does metalinguistic awareness relate to positive transfer from English?

Let us first take a critical look at the central variable to be determined regarding its predictive value in the transfer process, that is, metalinguistic awareness (MLA). Our primary aim was to isolate MLA from the other influential factors and to examine its role in the active use of English (L2) background vocabulary when French-speaking Quebeckers were presented with mostly unknown words in German (L3), which they were asked to translate into French (L1). Based on a large body of research on third and additional language development, of which MLA appears to be a driving force, we expected this "deliberately directed attention to the structural features of language and manipulation of its components" (our definition of MLA, see section 2.3.1.4) to be related to positive lexical transfer, regardless of proficiency and exposure levels in the source language, English. Our primary measure of MLA was a shortened version of the THAM-3 (Pinto & El Euch, 2015), which assessed a reflexive type of awareness, given that it focused on the participants' conscious manipulation of structural and semantic features of their native language, French. In addition, a secondary measure of MLA emerged from the analysis of the TAPs, which revealed distinct levels of "crosslinguistic awareness" (Jessner, 2008a), that is, the conscious manipulation of structural and semantic correspondences across languages. Given that this kind of reasoning reflected associations between lexical items from different languages by means of inferences based on metalinguistic processing, this secondary measure was taken to represent an applied dimension of MLA, which was added into the equation for the purpose of data triangulation.

As outlined in the previous chapter, positive lexical transfer was the response variable in two independent models of binomial logistic regressions, one of which included the results

of the THAM as the exclusive measure of MLA, whereas the other featured both of them. In both cases, MLA was the strongest predictor of positive transfer. The first analysis qualified MLA, as represented by the THAM scores only, as the primary influential factor (X^2 = 8.83, p = .003), followed by English language proficiency (X^2 = 6.62, p = .01), both of which were found to predict positive transfer significantly. These findings were consistent with our predictions. In fact, numerous researchers have underpinned the crucial role of MLA in the acquisition process of multilingual speakers in various areas of language, such as phonology (e.g., Marx & Mehlhorn, 2010), syntax (e.g., Bardel & Falk, 2007), and the lexicon (e.g., Dressler et al. 2011), as well as reading (e.g., Peyer et al., 2010) and writing skills (e.g., Cenoz & Gorter, 2011). However, MLA has not yet been investigated among the variables that could predict such positive transfer between non-native languages. Still, the research community appears to agree on the fact that heightened levels of MLA should be responsible for specific processing mechanisms of speakers and learners of third and additional languages, namely by enabling them to make fruitful associations with any of their previously acquired languages (e.g., Jessner, 2006; Singleton & Aronin, 2007). This apparent ease with which speakers of multiple languages adapt to different speech situations by making use of their cognitive and social flexibility to choose the appropriate linguistic elements to produce or infer meaning, has also been referred to as "multicompetence" (Cook, 1992, 2002, 2007) or "multilingual proficiency" (Herdina & Jessner, 2002). In the light of these appraisals of MLA as a major asset for further language learning, the present study aimed at measuring the contribution of this variable to the process of positive transfer from L2 to L3, along with the most discussed influential factors for such learning contexts, i.e., typology, proficiency and recency (see above).

The protocol that was chosen in the present research allowed for a measurement of MLA independently from proficiency and exposure measures, namely by using the THAM as the primary instrument to operationalize MLA. Finally, the results of the first regression analysis revealed the prominent role of this variable in the felicitous use of English L2 background vocabulary to understand new words in a related L3, German. This finding suggests that a reflexive kind of awareness, such as evidenced by higher scores on the THAM, enabled our participants to make appropriate form-meaning connections based on

their L2 English, and that this aspect was more important than their level of L2 proficiency. Furthermore, the qualitative analysis of the TAPs revealed some specific mechanisms that were used to justify assumptions based on crosslinguistic similarities. For instance, the search for semantic or syntactic cues in the accompanying sentences or the attempt to establish links with other learning events were most often associated with appropriate assumptions resulting in positive transfer. Hence, even if it was possible to resort to positive transfer from English on relatively low levels of crosslinguistic awareness, i.e., what we called *Intuition* and *Noticing*, respectively, most of the participants who analyzed the target items and sentences more thoroughly, namely at the levels of *Meta-Awareness* or even *Underlying Rule*, and in a larger number of cases, also tended to come to the right conclusions.

What can be retained from this observation is that MLA also seems to play a role in very basic learning situations such as the first encounter of isolated lexical items in the target language. In fact, even if the degree of typological similarity between the items (or structures) of the given source and target language predetermines part of the associations that could possibly be made by learners, our participants solved the task by means of a variety of approaches, as evidenced by the think-aloud data. Namely, some seemed to ignore the sentences and only relied on phonological representations, mostly by means of clang associations between the target item and any other words that came to their minds. Others tended to analyze orthographic features in relation to sound, and yet others tried to find lexical or syntactic cues in the sentences. Given the typological similarity between the source and the target language[61], any of these strategies, even the most intuitive, could have led to correct translations of the German test items based on their English counterpart. However, the deeper the analysis of the items themselves or of co-occurring words and structures in the sentences, the better our participants managed to come to the right conclusions. That is, their levels of crosslinguistic awareness, as evidenced by the MLA measure of the TAPs, really did seem to impact their transfer rates. This was illustrated by the relative distribution of both raw and normalized MLA scores from the TAPs in relation

---

[61] As outlined with respect to the Methodology (Chapter 4), only the 28 test items that had typologically related counterparts in English were included in the measure of positive transfer.

to raw and normalized transfer scores in Figures 15 and 28, respectively.

Moreover, the second regression model amplified the role of MLA in this transfer process. In fact, this secondary measure, which reflected an applied awareness, namely with respect to crosslinguistic correspondences, reached the highest statistic in the model ($X^2 = 26.44$, $p < .001$), followed by the MLA score from the THAM ($X^2 = 4.55$, $p = .03$). However, it remains crucial to consider the conceptual limitations regarding the operationalization of this secondary measure of MLA. Apart from the overlap of the zero-ratings between the measures of positive transfer and the MLA score from the TAPs, it must also be born in mind that for transferred items, the minimal MLA score was of 1 (*Intuition*), but most often equaled 2 (*Noticing*). Consequently, the second regression analysis may still be considered a distorted picture of the relationship between MLA and positive transfer, given that the two measures were based on the same data set. This is precisely the reason why we ran two independent regression analyses, both of which were examined for predictive power and model fit (see section 5.5.3).

Finally, even if the second model had a slightly higher concordance index, thus indicating that it could account for a larger part of the variance, it must be acknowledged that the area under ROC curve represented approximately 60 % in both cases, thus indicating only a very small difference with respect to their predictive power. As previously mentioned, models featuring indices below 0.7 could be rejected in clinical health research. However, it is important to consider that the physiological effects of pharmaceutical products can and should be predicted very accurately, whereas it is far more difficult to explain human behavior on the basis of specific factors. In turn, both of our models reflect a certain degree of predictive power with respect to the response variable. Namely, if we were to draw random pairs of observations, the predictions made by either model would be true in roughly 60 % of the cases. What should be retained from these analyses is that the first regression model, featuring the THAM as the exclusive measure of MLA, could by itself account for considerable part of the variance. In other words, if the MLA score from the TAPs was not entered into the equation, the model still served as an explanation of the response variable. Moreover, this first regression model was also consistent with data, as

shown by a Hosmer and Lemeshow goodness-of-fit test (see section 5.5.3).

With regard to this second and central research question, we may conclude that MLA is in fact related to positive transfer and that this relationship is independent of the respective levels of L2 proficiency or exposure. More precisely, it was shown by collinearity statistics that none of the predictive variables could be linearly predicted from each other, as indicated by variance inflation factors between 1 and 2 (see section 5.5.1). Our second hypothesis stating that MLA would predict positive transfer from English (see section 3.5) was thus confirmed. An elaboration on the kinds of conclusions that can be drawn from our study regarding the concept of MLA and its role in the L3 learning process will follow in section 6.2.

6.1.3 Does English proficiency relate to positive transfer from English?

In line with the extensive discussion on influential factors affecting transfer phenomena in Chapter 2 as well as the literature review of studies on positive transfer having focused on either of these factors in Chapter 3, source language proficiency was also shown to affect transfer rates in the present study. More precisely, the respective scores that our participants had attained on the Michigan Test of English Language Proficiency (MTELP) were shown to predict the extent to which these learners resorted to English (L2) background vocabulary in order to make sense of the German test items that they were asked to translate into French (L1). In fact, the statistical test results revealed that English language proficiency, as assessed by the MTELP, was the second most influential factor to predict transfer rates, when the THAM-3 was the only measure of MLA. In other words, the logistic regression model that did not include the metalinguistic ratings of the TAPs as a distinct variable, qualified English language proficiency as a predictor of positive transfer. Precisely, a Chi-square statistic of 6.62 and a p-value of .010 pointed to a significant relationship between L2 proficiency and positive transfer from the L2. As stated in the previous chapter, when the MLA measure of the TAPs was entered into the equation in the second regression model, this relationship was not significant anymore ($X^2 = 3.12$, $p = .08$). On the one hand, these findings suggest that proficiency in English is positively related to transfer rates. On the other, the importance that has typically been assigned to this variable

in a variety of transfer studies (e.g., Lindqvist, 2010; Odlin & Jarvis, 2004) may have to be reconsidered in relation to the learners' levels of metalinguistic development, which might affect the recognition of related background vocabulary to an even greater extent than and independently from L2 proficiency. Not only was MLA a stronger predictor of positive transfer in this study, but the specification of the different dimensions of MLA also strengthened its predictive value, while weakening that of L2 proficiency.

In other words, the protocol that was adopted in the present study did reveal a relationship between L2 proficiency and positive transfer from L2 to L3. This was shown both in the regression and in the correlational analyses, thus corroborating previous research on transfer between typologically related non-native languages. However, the protocol also revealed that the cognitive dimension of multiple language acquisition might have been underestimated in the study of language transfer. Namely, as discussed with respect to the first research question, the applied dimension of MLA was a stronger predictor of positive transfer than L2 proficiency. In sum, our third hypothesis stating that English language proficiency would predict positive transfer from English (see section 3.5) was only partly confirmed. When the applied dimension of MLA was entered into the equation, English proficiency did not qualify as predictor anymore, even though it was positively correlated with the transfer measure.

In the following subsection, the role of lexical access in multilingual lexicons will be further discussed with respect to the results relative to L2 exposure.

6.1.4 Does exposure to English relate to positive transfer from English?

Another aspect that was discussed in the literature on transfer between non-native languages was that of *recency* of use or exposure (e.g., Williams & Hammarberg, 1998). Namely, along with typology and proficiency in the given source language, the so-called *recency* factor is supposed to impact the level of activation of words and structures from a non-target language (e.g., Green, 1986, 1998). In the present study, this aspect was investigated with respect to frequency of L2 use, on the one hand, and L2 exposure, on the other. Interestingly, our findings regarding both these variables did not indicate a

relationship with the response variable. This does not reflect previous research findings such as Dewaele's (2001), where the regularity of L2 use did impact the frequency with which participants produced lexical interventions based on that L2. Rather, in line with Williams and Hammarberg's (1998) who had coined the notion of *recency*, words from a recently and/or frequently used language did appear to compete in the selection process in multilingual production in Dewaele's study. Similarly, Bayona (2009) found that the amount of current exposure to French[62], along with the fact that her participants interacted with friends and family and that they read in French was significantly correlated with the amount of French-based transfer in her study (Bayona, 2009, pp. 84-88). Moreover, according to findings reported by Tremblay (2006), the amount of L2 exposure may even be expected to impact transfer rates to a greater extent than L2 proficiency (Tremblay, 2006 pp. 116-117), which was clearly not the case in the present study. In both regression models, neither of the two exposure measures, that is (1) the frequency of English language use in different non-instructional contexts and (2) the amount of time spent in English-speaking environments, was related to positive transfer from English.

In fact, there might be different ways of interpreting these divergent results. First of all, the studies reported above investigated crosslinguistic influence with a focus on what may be called deviant TL production data. Clearly, the fact that our study was concerned with the processing of new L3 vocabulary presented in a controlled setting provided a different context for possible association patterns than the more or less spontaneous production of speech or text in an additional language. More precisely, most of the instances of (negative) transfer reported in the other studies reflected unintentional language switches, whereas the kind of (positive) transfer observed in our study involved the conscious search of related background vocabulary. If we consider these two kinds of multilingual processing with regard to lexical access, it appears that the activation patterns suggested by Green (1986, 1998) only apply to contexts that call for the use of a single language. According to Green, "speech production involves controlling the activation of internal representation of words" (Green, 1986, p. 215). In other words, if only one language is selected, the other

---

[62] Participants were asked to indicate the percentage of time during which they were currently exposed to French.

language(s) that might have reached an equal or higher level of activation in the speaker's mind must be inhibited. As outlined in the conceptual framework, the notion of "inhibitory control" was taken up in de Bot's (2004) model of the multilingual lexicon, where relative activation levels are largely determined by the amount of contact and use (see de Bot, 2004, p. 10). Research findings like the ones reported above (Bayona, 2009; Dewaele, 2001; Tremblay, 2006) could substantiate these assumptions by establishing a relationship between L2 exposure and/or frequency of use and learner errors based on interference from that language. However, in contexts where the co-activation of words from another language is desirable for a learner to make successful assumptions about the TL, the links that are established between the unknown target and potentially helpful L2 material might not depend as much on a certain activation threshold of the background language but rather on the learner's specific approach to solving such comprehension problems. In other words, as outlined in the conceptual framework, it is not because the appropriate source language material is accessible that "the learner will do the necessary looking or come to the right conclusion about just how congruent a cross-linguistic correspondence is" (Odlin, 2003, p. 443).

A second consideration that seems crucial for the interpretation of our results is concerned with the kind of language exposure that might have impacted the active use of English cognate vocabulary in the present study. As discussed in our presentation of the results, different aspects of L2 exposure were reflected in the data, to account for a broad definition of the term, including receptive, productive and interactive language activities. For our primary exposure measure labeled "Frequency of use", a principle component analysis revealed distinct loadings for three groups of activities in the self-evaluation grid, subsumed under (1) oral interaction and listening, (2) written interaction and writing and (3) interaction with a machine. Interestingly, the oral component of English language use was the only one that could not be related to our measure of English proficiency, while the strongest correlation could be established with the written component. As for the basic exposure measure where participants had to indicate the relative amounts of time spent in different English-speaking environments, no relationship to proficiency was found. At the same time, this secondary exposure measure was moderately and significantly correlated

with the oral component of English language use. In turn, the fact that this basic exposure type did not relate to our proficiency measure could possibly be explained by the fact that the linguistic environments in question provided more opportunities for oral than written language activities. By extension, a more in-depth analysis of written component of L2 exposure and use might in fact reveal an indirect relationship to the type of positive transfer that was the object of our study. This assumption is motivated by the following observations: First, high levels of MLA appeared to be helpful to confirm translation choices based on positive transfer from English (see following section). Second, according to research on literacy development in both L1 and L2 (see section 3.3.2), metalinguistic thinking appears to develop especially through language activities based on reading and writing (e.g., Gombert, 1990). Thus, L2 activities involving the written code – be it for receptive, productive or interactive purposes – may be expected to impact MLA. As a consequence, if the amount of literacy-based activities in multiple languages significantly predicts MLA, this particular type of L2 exposure might also be related to positive transfer, with MLA as the mediating variable. Such a finding would extend recent research suggesting that the positive effects of L1 and L2 literacy practices on the development of L3 reading skills are mediated by MLA (Rauch et al., 2012). Finally, it is possible that a larger data set of literacy-based activities in our participants' background languages might have revealed a relationship to transfer rates in our study.

In sum, L2 exposure such as it was operationalized in the present study was not shown to relate to positive transfer from English, thus rejecting our fourth hypothesis stating that L2 exposure would positively correlate in transfer rates (see section 3.5). This finding seemed inconsistent with previous research on lateral transfer, where the closely related concepts of recency and frequency of source language use could predict transfer rates. However, this discrepancy may be related to the different kinds of processing under observation, i.e., monolingual speech production in L3 where the influence from other languages is ideally inhibited versus the comprehension of new L3 items where the emergence of related L2 vocabulary is a welcome basis for lexical inferencing. Finally, we also considered the possibility that the absence of a relationship between L2 exposure and positive transfer was related to the global measure of exposure, whereas a more detailed description of L2

literacy practices may have revealed a relationship between this specific type of L2 exposure and positive lexical transfer.

6.1.5 Do levels of interest in the German language and culture, respectively, relate to positive transfer from English?

Our last research question was concerned with the affective dimension of language learning, regarding the respective levels of interest in the German language and culture. With reference to Hufeisen's (2000) Factor model (Figure 1, Chapter 2), it seemed indispensable to take at least one affective factor into consideration, given that the other predictive variables were related to cognitive and linguistic aspects of the learning process. As a matter of fact, the degree of motivation to learn a certain language and the attitudes toward the target language community were shown to affect the learning process of both SLA (e.g., Dörnyei, 1997; Gardner, 1985) and TLA (e.g., Hufeisen & Gibson, 2003). However, the complexity of each of these concepts calls for instruments of data collection that are generally rather voluminous questionnaires in themselves (e.g., AMTB, Gardner, 1985). However, as mentioned above, the context of the present study did not allow to examine this variable accordingly, which led us to include two questions in the background questionnaire. More precisely, participants were asked to rate their levels of interest in the German language (Q19) and culture (Q20), respectively, by referring to a Likert-scale, ranging from 'très faible' (1) to 'très fort' (5). As briefly discussed in the previous chapter, neither of the two interest variables was positively related to the participants' transfer rates. Moreover, both the regression models presented in section 5.5 of the previous chapter suggested an inverse relationship for 'Interest in the German culture. Only in the second regression model, this relationship was significant ($X^2 = 4.18$, $p = .04$), even if the p-value was only slightly below the 5 % threshold.

Aiming at a plausible explanation for this finding, two interpretations were taken into consideration. First, it seemed important to consider the difference between the two interest variables and to ask oneself why higher levels of interest in the target language culture could have a negative impact on the participants' transfer rates, whereas interest in the target language did not affect the response variable at all. The only explanation that seemed

plausible in this respect was the following: If certain learners decided to learn the target language merely out of interest for cultural aspects of target language community, this may have been cross-related to a lack of interest in the formal and functional aspects of the language. In other words, if higher levels of interest in the German culture[63] were in fact related to a lack of motivation for instructional language learning, including activities such as the word translation task, this might result in lower transfer rates. Of course, this attempt for an explanation is highly speculative and it would have to be corroborated by subsequent investigations into the target population. Unfortunately, however, we were unable to get access to the participants after the data collection phase for further inquiries. Notwithstanding, this observation suggests that a more thorough investigation of affective variables might add to our understanding of the transfer process with respect to individual differences. In fact, this is also the starting point for our second interpretation of this supposedly inverse relationship between 'Interest in the German culture' and positive lexical transfer from English. Namely, by providing a five-point Likert-scale to indicate levels of interest in the German language and culture, attitudinal and motivational factors were clearly oversimplified. In turn, we have to consider the possibility that the questions were not constructed in a way that could provide valid data of the participants' levels of interest in the target language and culture. Consequently, future investigations into positive transfer between the non-native languages may benefit from using comprehensive test batteries such as the AMTB (Gardner, 1985) to investigate the potential relationship to the active use of background languages of multilingual speakers. After all, an exhaustive account of attitudinal and motivational variables may appropriately reflect the affective dimension of Hufeisen's (2000) Factor Model (see Figure 1) in relation to the L3 learning process. Based on the findings of the present study, however, our last hypothesis stating that levels of interest in the German language and culture would relate to positive lexical transfer from English (see section 3.5), was rejected.

To sum up this section, we may conclude that the most prominent factor that was shown to influence positive transfer in the present study was in fact MLA, followed by English

---

[63] In the questionnaire, "Culture" was specified by the following aspects: customs, history, gastronomy, arts, architecture, literature, etc. (see Appendix A)

language proficiency, whereas none of the remaining predictive variables in our protocol appeared to be related to the response variable. It was the THAM, on the one hand, and the MLA score from the TAPs, on the other, that were identified as the strongest predictors of positive lexical transfer from English in two independent regression models, followed by L2 proficiency. Moreover, the methodological approach to the concept of MLA allowed us determine two aspects of awareness – a reflexive and an applied dimension – which seem to be closely related to each other, while contributing to the comprehension of unknown lexical items in L3 on the basis of crosslinguistic correspondences. The implications that these observations might have on future research, both with respect to conceptual and methodological considerations, will be discussed in the following section (6.2).

6.2   The role of MLA in the L3 learning process

Having identified MLA as the strongest predictor of positive lexical transfer, the goal of this section is to provide a crtitical analysis of the scope of our investigation, namely regarding the concept of MLA itself, and its role in the learning process of a third or additional language. More precisely, we will discuss what evidence can be drawn from our observations with reference to our central methodological choices. In the first place, we will summarize what our protocol revealed about the concept of MLA.

As discussed in the conceptual framework, we adopted a psycholinguistic approach to the concept of MLA, in which the notion of *awareness* was considered in terms of a mental activity (instead of a latent state) that involves a minimal amount of attention on the part of the learner, i.e., the detection of a stimulus (see Schmidt, 1994). Thus taking into consideration the level of *noticing*, "awareness is minimally some sort of cognitive or behavioral change taking place during input processing without any direct association with or reference to the underlying grammatical rule" (Leow, 2015, p. 47). What is of crucial importance in our definition of awareness is that the notion of *explicit learning* (a process) is not directly related to *explicit knowledge* (a product). In other words, by adopting Schmidt's conception of explicit learning, the initital stage of which is "noticing" or "focal awareness" (see Schmidt, 1990, p. 132), then this process can, but need not, result in explicit knowledge (see also section 2.3.1.2). Similarly, the notion of *metalinguistic* that is

commonly associated with the explicit reference to grammatical categories or "metalinguistic terms" (e.g., Alderson et al., 1997) was considered at its very conceptual basis. By contrast to the notion of *epilinguistic,* which refers to linguistic behaviour that is not controlled, i.e., where attention is not actively focused (e.g., Gombert, 1990), we identified the fundamental nature of *metalinguistic* activities in terms of intentional control. In other words, the distinction between the two concepts is based on the intentional dimension of consciousness, rather than on depth of analysis (see section 2.3.1.3).

With reference to Bialystok (2001a) and Jessner (2006), someone who is metalinguistically aware was broadly defined as having the ability to deliberately direct his or her attention to the structural features of language, and thus to manipulate its components. Keeping in mind that this punctual awareness did not necessarily involve the explicit knowledge of specialized terminology or underlying rules, we adopted a fine-grained methodological approach to be able to identify levels of MLA. This included instances where learners exhibit a sort of "intuitive awareness" (Tunmer & Herriman, 1984, p. 19) thus reflecting the onset of noticing, an initial stage of awareness, while higher levels were reflected by varying degrees of explicitness, i.e., the extent to which linguistic features are being analyzed and manipulated. For example, the recognition of certain patterns or regularities may qualify as an intermediate stage, whereas the explicit naming of grammatical categories or underlying rules were taken to reflect the highest level. Our protocol enabled us to differentiate between different levels of MLA by means of two independent measures. Our primary instrument, the adapted version of the THAM-3, featured 16 test items, each of which required a linguistic (L) and a metalinguistic (ML) answer relative to different semantic, pragmatic and grammatical aspects of the participants' native language French. According to our own interpretation and in contrast to the authors of the test (Pinto et al., 1999; Pinto et El Euch, 2015), the L score of the test was thought to reflect an initial stage of awareness, that is, an intuitive dimension of MLA (see e.g., Leow, 2000; Schmidt, 1990). The ML score, on the other hand, represented two levels of more or less analyzed (explicit) knowledge, the highest level of which mostly involved the extrapolation of underlying rules or patterns (see section 4.3.4.2).

Given that we also had access to introspective data from a translation task that involved the manipulation of our participants' target language German (L3), our secondary measure of MLA was based on the qualitative analysis of the TAPs. Here, we differentiated between four levels of awareness, that were labeled *Intuition* (I), *Noticing* (N), *Meta-Awareness* (MA) and *Underlying rule* (UR). Whereas I and N were taken to represent the level of *noticing*, MA and UR referred to what Schmidt had differentiated as *understanding*. Even though it may seem out of place to justify two levels of noticing, the first category (I) was added to our coding system in order to distinguish between verbalizations where an intuitive correspondence to a typologically similar cognate was only implicit in the verbalization (I), as opposed to instances where the source of knowledge was explicitly stated (N), but not justified. Apart from this first label (I), the three other categories were inspired by Leow (1997) and other researchers (e.g., Sachs & Suh, 2007) who made a subtle distinction between *noticing*, as evidenced by a cognitive change in the learner's verbal behavior, and two levels of *understanding*, the first of which (MA) referred to a recognition of a pattern while not including explicit references to underlying rules or other kinds of highly analyzed knowledge. In our protocol, this category was reserved for instances where participants explicitly analyzed correpondences in sound or spelling that inspired a given translation equivalent, or else when they used co-occurring words or structures in the supporter sentences to justify their answers (see section 5.2.2). Finally, only explicit references to underlying rules were rated at the highest level of awareness. In our protocol, verbalizations in this category (UR) occurred very rarely, given that no morphological or syntactic knowledge was necessary to translate the unknown test items from German (L3) into French (L1). In order to account for a wider array of metalinguistic processes, future investigations resorting to tasks that call for such deeper levels of analysis would be desirable, namely to present learners with the challenge to make use of their metalinguistic abilities. This aspect will be further discussed in the section on the limitations of the present study (see section 6.3) by taking a critical look at the suggested improvements as well.

Finally, in addition to a fine-grained analysis of MLA in terms of distinct levels of awareness, our protocol also allowed for a differentiation between two dimensions of

awareness, one that was labeled *reflexive* and the other *applied*. The distinction between these types was determined by the object of inquiry of the respective instrument and the kinds of metalinguistic resources required by each of the tasks. On the one hand, the THAM was designed to trigger the participants' reflection about their native language. Moreover, the manipulations that they were asked to perform on the linguistic features targeted by the different parts of the test were of descriptive nature. This is why the results of the THAM were taken to represent a reflexive type of awareness. On the other hand, the metalinguistic processes exhibited in the TAPs were thought to reflect an applied dimension of MLA, in so far as participants mobilized their attentional resources and their metalinguistic abilities to make sense of unknown words and structures in a third or additional language. Form-meaning correspondences had to be established between different languages and focal attention towards specific lexical and grammatical features was conducive to justifying translation choices based on strictly formal aspects of target and source language. More specifically, lexical inferencing was achieved by means of more or less elaborate analyses, which were taken to represent the degree to which participants applied their metalinguistic abilities to perform a search of potentially helpful cues across languages.

According to our conception of the cognitive processes that learners of multiple languages engage in, the two dimensions of MLA are closely related, in so far as they are believed to condition each other. If we assume that all language users "learn the probability of an interpretation given a formal cue in a particular context" (N. C. Ellis, 2006, p. 8), the focus of our inquiry was on the interpretation. Participants had to figure out if a given cue was appropriate. When interpretations were based on more thorough analyses of the target structure in relation to the immediate context (the supporter sentences) and/or other individual learning experiences, appropriate form-meaning mappings were generally more frequent. Moreover, higher levels of this applied dimension of awareness, as exhibited in the TAPs, were associated with higher scores on the THAM. This finding corroborates Jessner's (2008a) claim regarding the role of MLA in multilingual processing:

> Crosslinguistic awareness in L3 production can be defined as the awareness (tacit and explicit) of the interaction between the languages in a multilingual's mind;

metalinguistic awareness adds to this by making objectification possible. (Jessner, 2008a, p. 279)

In turn, the multiplication of experiences in which learners apply their metalinguistic abilities to justify new associations in their lexical network, are also expected to impact the reflexive dimension of awareness. These observations are directly related to the question of what the present study reveals about the role of MLA in the L3 learning process. As previously stated, many studies in SLA and especially in TLA have stressed the importance of MLA in the learning process of subsequent languages (see Jessner, 2008b for an overview). In fact, the advantages that learners/users of multiple languages seem to have over monolinguals and novice learners of a first non-native language may be related to higher levels of MLA. In the TLA literature, this relationship has been discussed under the label of "additive multilingualism" (Cenoz, 2003a).

In the present study, the role of MLA was investigated in a very specific context of L3 learning, that is, positive lexical transfer from a typologically related L2 when the target items were largely unknown to the participants. For various reasons that will be taken up with respect to the limitations of our study (see section 6.3), our response variable was operationalized by a word translation task, which it was possible to solve at an intuitive level of awareness. At the same time, higher levels of MLA, in terms of explicit manipulations of more or less analyzed knowledge, did have a positive impact on transfer rates. This finding suggests that the active manipulation of formal and structural aspects of the TL in relation to the learners' background languages (especially typologically related ones) is an asset, even for very basic processes such as cognate recognition. The applied dimension of awareness that was measured by means of the TAPs was the strongest predictor of positive transfer in our study. From the statistical relationship that was shown to exist between the reflexive and the applied dimensions of awareness, we assume that the kind of reasoning that learners use to analyze structures in their L1 (reflexive awareness) are also applied when they are engaged in learning events that involve the active manipulation of their non-native languages (applied awareness). In short, learners get used to focusing their attention on structural features of language in general. In line with the above discussed notions of *common underlying proficiency* (CUP) (Cummins, 1991, 2005),

multicompetence (Cook, 1992, 2002, 2007) as well as *multilingual proficiency* (Herdina &
Jessner, 2002; Jessner, 2006, 2008a), our observations suggest that this capacity develops
with the multiplication and complexification of the learning events learners engage in.
Moreover, our interpretation of the reciprocal relationship between the two dimensions of
MLA is consistent with findings on literacy development (Durgunoğlu, 2002; Rauch et al.,
2012; Schwartz et al., 2007), which suggest that the positive impact of literacy skills in a
given source language (L1, L2, Lx…) on a variety of literacy-based activities in an
additional language can best be explained in terms of strategy transfer.

In sum, learners who are used to manipulating their languages in a variety of contexts
become more aware of the formal and functional aspects of language. Additonal languages
complexify these manipulations, because they have certain features that are similar and
others that are different from previously acquired structures. When learners get the chance
to become aware of these nuances, which is most easily achieved in instructional contexts,
they might come to understand the underlying rules that govern one or the other feature,
and become more proficient with respect to their analytic skills. In turn, if we know that
MLA is positively related to more efficient processing of new TL material, then how can
this be applied to the foreign language classroom? In this regard, the learning conditions
that favor the development of MLA in multilingual contexts should be identified. After an
extensive discussion of the limitations of our study in the following section, we will address
the pedagogical implications of our findings, in line with current research on multilingual
education.

6.3   Limitations

Most of the limitations of the present study are concerned with methodological choices and
conceptual issues that result from them. We will present and discuss the pitfalls that were
engendered by our methodology. At the end of each subsection devoted to one of the
selected instruments, we suggest possible solutions for the given problem.

6.3.1   Word translation task

There are two major considerations concerning the word translation task that was chosen to

collect data on positive lexical transfer. On the one hand, it may be regarded as too simple for the particular purpose of our research agenda, whereas on the other hand, it may be considered too complex for a variety of different reasons. First of all, a critical look at the choice of the test items reveals that the degree of similarity between the different target items and their English counterparts predetermined transferability, at least to a certain extent. Moreover, it did not necessitate a high degree of awareness to solve the problem. Thus, the verbalizations might not represent what participants were truly able to do in terms of thorough metalinguistic analyses. In other words, if we wanted to investigate the extent to which learners were actually able to manipulate TL words and structures by actively resorting to their background languages to solve comprehension problems, the task may in fact be considered too simple. However, we needed to construct a task that did not necessitate MLA to solve the problem, otherwise our argument would have been circular: *Does MLA predict a phenomenon that necessitates MLA?* This is why we did not opt for a more demanding task, such as text translation (e.g., Gibson & Hufeisen, 2003) or comprehension questions of an unknown language as in the Indonesian Language Test (Kuile et al., 2011).

Finally, the fact that prior knowledge could not be ruled out entirely by normalization made the transfer measure less robust. That is, participants who did not know any of the words were faced with the totality of the target items, for some of which it was easier to associate an English cognate than for others. Even though transfer rates were normalized by counting the mean scores for each participant depending on the number of unknown items that could have been transferred, the fact that prior knowledge was listed for different words across the sample, left some participants with a potentially more challenging task than others. Methodological approaches that might prevent this aspect from affecting the central measure in future investigations are exposed and briefly discussed in the subsection of this chapter pertaining to the selection of our participants (see 6.3.4). However, it should be pointed out that one of the shortcomings of the protocol was the absence of an initial questionnaire by means of which test items that were already known to some of the participants could have been eliminated prior to test administration.

6.3.2   Think-aloud procedure

As for the word translation task, there are also several considerations related to the think-aloud procedure that was used to collect data on the transfer process on the one hand, and on the levels of metalinguistic analysis, on the other. As discussed with regard to the methodology adopted for our study (section 4.3.3), the main conceptual challenge that researchers who use verbal protocols are faced with is related to what has been called reactivity (e.g., Bowles, 2010). In short, it is possible that the verbalizations themselves influence the test results by interfering with the learner's mental processes, as evidenced, for example, by slowing down reaction times, or by inducing more thorough analyses. In other words, the act of speaking their thoughts out loud may possibly enhance the learners' awareness during the testing, while it could also block certain inductive processes and thus lead to weaker performances (e.g., Bowles & Leow, 2005). This particular method was chosen with reference to several researchers who investigated the reactivity of TAPs in different SLA-based learning tasks, and who argued that concurrent probing did not tend to affect test results when the instructions did not prompt learners to attend to specific features encountered in the task (e.g., Leow & Morgan-Short, 2004). In the present study, the written instructions of the TAPs (see Appendix D) were non-metacognitive in nature, that is, participants were asked to verbalize their thoughts as they came, without elaborating on any particular feature. However, a specification was added to these instructions after the researcher noticed during the first testing session that some participants failed to elaborate on certain items that were already known (see section 4.4.1.1). Throughout this session, participants who provided translations without elaborating on how they came to this conclusion were prompted to indicate their source of knowledge. In the following testing sessions, the verbal instructions given prior to the task were adapted to explicitly incite participants to state the knowledge source of their translations. This additional instruction may in fact qualify this type of verbal report as metacognitive, thus increasing the potential reactivity of the TAPs. As previously stated, metacognitive TAPs were shown to be reactive in some receptive (Bowles & Leow, 2005) and productive tasks (Bowles, 2008). Yet what these researchers found were detrimental (instead of facilitative) effects on performance when the verbal report required elaboration on specific features of the targeted structures. At the same time, it was also argued with regard to problem-solving tasks with

different degrees of complexity that metacognitive verbal reports only seem to interfere with task performance when the given task already puts a high cognitive load on the participant (see Bowles, 2010). As previously stated regarding the low level of complexity of the work translation task, it was improbable that the TAPs would have detrimental effects on task performance in the present study. In turn, it may be assumed that if the TAPs had an effect on the participants' cognitive processes while solving the task, it would most likely have been facilitative. With reference to Bowles' (2010) observations concerning the few studies where facilitative but only small effects were found (see section 4.3.3.1), it was assumed that potential effects would not invalidate the findings of the present study. However, it must be acknowledged that the act of verbalizing may have helped participants to focus their attention on various formal and semantic aspects of the target items in relation to their background languages. In this regard, future investigations resorting to the use of TAPs should consider the inclusion of a silent control group (see Bowles, 2008; Bowles & Leow, 2005). Unfortunately, this was beyond the scope of the present study, but will have to be considered in future investigations using this methodology. Moreover, follow-up interviews on the think-aloud procedure or other retrospective measures in combination with the TAPs may provide crucial insights into the extent to which verbalizations shape learners' thoughts, instead of merely reflecting them. For example, Swain (2006b) reports on research having approached verbalization through the lens of sociocultural theory, which strongly suggests that speech functions as a motor of thought (e.g., Smagorinsky, 1998) thus challenging the information processing approach adopted by Ericsson and Simon (1993) who claimed that verbal reports reliably represent cognitive processes, as long as instructions do not put specific demands on the learner (see section 4.3.3.1). As pointed out by Swain:

> First, the process of verbalization itself transforms thought, drawing attention to some aspects of the environment and not others, solidifying meaning, and creating an observable artifact. Secondly, as an observable artifact, it can be reflected upon, questioned, manipulated and restructured. And thirdly, internalization of this now differently understood externalized artifact may occur. What this implies is that verbal protocols not only potentially transform thinking, focussing it in highly specific ways, but also are the sources of changes in cognition. In other words,

speech mediates learning and development. (Swain, 2006b, p. 101)

What should be retained from the above discussion is, on the one hand, that the use of TAPs as a research methodology remains questionable in terms of reactivity, which future studies could control for by adding a control group who complete a task without being asked to verbalize. On the other hand, the specific role of verbalization in explicit learning should be investigated. Namely, if we assume that the act of speaking shapes and refines mental processes by making knowledge accessible, TAPs and other types of verbal data seem to be promising research methodologies to explore consciousness-raising activities in instructed SLA and TLA. In turn, the obvious challenge for researchers using verbal reports is that of extensive data triangulation (see also section 6.4).

Apart from the fact that the TAPs may have been conducive to making associations between languages that led to correct translations, another aspect of this data collection method must be kept in mind, with regard to the levels of metalinguistic analysis that were identified on the basis of these protocols. First, absence of verbalization cannot necessarily be taken to mean absence of analytic reasoning. In other words, those who verbalize justifications based on structural features, form-meaning correspondences, or even underlying rules, do not necessarily carry out more analyzed mental processes than those who do not verbalize their ways of reasoning (see section 5.2.2.1). In turn, if we concede that good analyzers are not necessarily good verbalizers, this would imply that the levels of MLA such as they were determined by the qualitative analysis of the TAPs do not reliably reflect the participants' actual degrees of awareness. However, our research agenda is based on the conceptual consideration that what defines "meta-processes" in language learning is that they are accessible to consciousness, and thus verbalizable (see Karmiloff-Smith, 1986). A more commonly used term in SLA is that of *explicit learning*, namely referring to "a process during which participants acquire conscious (explicit) knowledge" (Rebuschat, 2015, p. XIII). On the other hand, what is typically referred to as *implicit learning* is thought to result in an "intuitive form of knowledge that goes beyond what can be verbalized" (Schmidt, 2010, p. 726). It follows from these basic definitions that verbalization plays a central role for the operationalization of *explicit* vs. *implicit* learning processes. This is reflected in the levels of awareness that were identified on the basis of

verbal data in the present study. Namely, according to previous studies where distinct levels of awareness were analyzed in TAPs (see section 4.5.1.1), our taxonomy depended on the extent to which participants verbalized their thoughts. It remains unclear whether the less elaborate verbalizations or the mere absence thereof can in fact be taken to reflect absence of awareness. This methodological shortcoming has been addressed in a recent contribution by Rebuschat and his colleagues (Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015) who suggest data triangulation to tease apart the different explicit and implicit processes that learners engage in during a given task. In addition to concurrent and retrospective verbal data, these researchers also used what they called "subjective measures of awareness" in the forms of "confidence ratings" of how sure participants were about their own answers and "source attributions", where they indicated whether their answers were based on *guess*, *intuition*, *memory* or *rule knowledge* (Rebuschat et al., 2015, p. 309). According to the authors, these measures enabled them to differentiate *explicit* from *implicit* processes of knowledge construction and thus to corroborate acquisition theories that posit the possibility of co-existence of both kinds of processes in particular learning events (e.g., Hulstijn, 2015a). What can be retained from these findings with respect to our own study is that data triangulation of awareness measures can be pushed further to include at least one instrument that does not elicit verbal data. In other words, while the THAM and the TAPs enabled us to assess two dimensions of awareness, it would have been helpful to use a subsequent measure to corroborate observations based on the verbal reports that were collected during the translation task.

Finally, apart from the conceptual considerations regarding the think-aloud procedure, it should also be recalled that there were some inconsistencies with respect to coding in terms of levels of metalinguistic analysis. Not only do we have to acknowledge that the categories labelled *Intuition*, *Noticing*, *Meta-Awareness* and *Underlying rule* (see section 4.5.1.1) would benefit from more clear-cut boundaries, but also the fact that the second rater did not have sufficient knowledge of German must be considered a weakness regarding the analysis of these highly variable qualitative data. As previously mentioned (see section 5.2.2.3), future investigations would need to rectify the selection of raters by recruiting candidates who have sufficient knowledge in all the manipulated languages, in order to

judge the internal logic (or lack thereof) exhibited in the verbal protocols.

6.3.3   Adaptation of the THAM

As for the metalinguistic ability test that was adapted from its original version, one conceptual consideration must be discussed. In order for the instrument to be consistent with the definition of MLA that was adopted in the present study, it was judged appropriate to exclude the test items that did not feature a linguistic (L) dimension. According to the authors of the test, this dimension did not represent strictly metalinguistic processes, given that intuitive awareness was sufficient to answer these questions (Pinto et al., 1999). As indicated throughout the present research, we adopted a larger definition of MLA, in which initial stages of awareness were also taken into consideration, given that they represent the basis upon which more analyzed knowledge is being built. In line with Schmidt's (1990, 1994, 2010) work on the levels of *noticing* and *understanding*, and with Leow's investigations into levels of awareness, we adopted finer-grained distinctions within each category (see previous section). Namely, what has been labeled "low" (Leow, 2000) or "intuitive" (Tunmer & Herriman, 1984) awareness is taken to represent an initial stage of the metalinguistic process. In other words, the specific process that learners must engage in to answer the L answers is thought to occur at the level of *noticing*. Moreover, if the term *metalinguistic* refers to processes where (more or less analyzed) knowledge is applied consciously and deliberately, then error correction does in fact qualify as a metalinguistic phenomenon. Based on these conceptual nuances, the L dimension of the THAM was included in the final MLA score.

Finally, even if all the sections of the original test were represented in our adapted version that included only test items in which both L and ML answers were required, we must concede to a conceptual limitation that resulted from this methodological choice. In fact, our adaptation affected the relative distribution of numbers of test items per section. While the *Comprehension* and *Figurative language* sections were reduced by little more than half their size, the *Acceptability* section remained intact and thus represented a much larger proportion of the totality of the retained test items. In turn, given that the acceptability part was too long compared to the other parts of the test, it would have been advisable to

counterbalance the investigation by reducing this section as well. That is, in order to prevent a bias towards metagrammatical analyses, we would have needed a larger proportion of metasemantic questions to tackle this conceptual issue.

6.3.4 Participants

Several aspects of the sampling procedure have to be taken into consideration with respect to the limitations of the present study. First, for an investigation into positive transfer where distinct levels of MLA and of L2 proficiency are expected to impact the response variable to varying degrees, the group of students who took part in this study may be considered as too homogenous. It may have been advisable to test groups of learners who are not exclusively enrolled in a language program in order to investigate more thoroughly whether and in what ways considerably lower levels of MLA and L2 proficiency, as would be expected for other populations with less language learning experience, would affect the strength of the predictive variables. Whereas the relative homogeneity of our sample was clearly an advantage regarding the selection of learners with similar language backgrounds, future investigations may want to aim for a wider array of MLA and proficiency levels in order to assess the relative contribution of each of these factors more thoroughly.

A second issue that should be kept in mind for future investigations into positive transfer when this concerns the manipulation of novel items or structures in the target language is the issue of prior knowledge. As previously mentioned with respect to the specific coding procedure of the transfer measure, it was necessary to normalize rates for each participant depending on the number of test items that were already known. Even if this calculation allowed us to quantify the final scores accordingly, the fact that test items had different degrees of difficulty affected these normalized scores also qualitatively. In sum, it would have been desirable to exclude the possibility of prior knowledge altogether. The only way of ruling out this factor after the task had been administered would have required discarding all test items of which at least one participant had prior knowledge, which would have left us with only ten (10) out of twenty-eight (28) test items. Such a large decrease in the number of observations may have diminished the potential of our protocol to account for the targeted phenomenon. In turn, if the goal was to prevent prior knowledge across the

board, future inquiries resorting to word translation tasks of this type would have to aim for a much larger number of test items prior to investigation, so that a subsequent reduction in size would not invalidate the protocol.

At the same time, within a usage-based approach to language learning, the endeavor may be considered somewhat incongruous given that variability due to individual experiences with the target language are likely to surface in any learning situation. Especially in the highly complex lexical networks of multilingual speakers, associations between words and structures can emerge unexpectedly. According to the premise that "usage is rich in latent structure, and [that] learners apprehend this structure in the large part by means of implicit learning" (N. C. Ellis, 2015, p. 11), even novice learners who are supposed to know nothing about the target language might in fact have some intuitive knowledge about certain aspects of it that cannot be controlled for. Instead of ignoring this variability as noise, methodologies that allow for descriptions and analyses of these differences are thought to extend our understanding of transfer processes in multilingual contexts.

6.4   Implications for future research

The implications of the present study for research communities who investigate multilingual development are two-fold. On the one hand, our findings provide a number of conceptual and methodological insights into the study of MLA and its role in the acquisition process of third or additional languages. On the other hand, they might inform pedogical practices and the promotion of individual multilingualism in educational contexts. Both of these aspects will be addressed in this last part of the discussion. With reference to current research in cognitive psychology and applied linguistics, we will sum up our central observations regarding the learning process to finally expand the scope to possible repercussions on the multilingual classroom.

The central goal of the present study was to examine the degree to which different learner-related variables influenced positive lexical transfer from a typologically related L2, English, to a low proficiency L3, German. In line with our hypotheses, our findings revealed that MLA was the strongest predictor of positive transfer, followed by English

language proficiency. The particular focus of our study was on the process of explicit learning. The aim of our protocol was not to reveal insights into the specific product of learning, i.e., explicit knowledge. In other words, the present study was concerned with awareness as 'construction' instead of awareness as 'reconstruction' (see Leow et al., 2011). That is, we looked at online-processing of novel words and the kinds of connections (to previously acquired languages or to other context-related cues) that appeared to benefit from MLA, at a specific moment. Whether learners retain these connections, whether they are able to reproduce them in different usage events was not the focus of our research agenda. What our protocol revealed was that the process of becoming aware of correspondences between unknown target items and related background vocabulary benefits from the conscious manipulation of the target structures. The more explicit these manipulations, the better the chances of appropriate form-meaning mappings. Even if the persistence of these representations in the learners' minds cannot be directly extrapolated from our results, it appears that explicit processes are in fact likely to result in explicit knowledge: "Our consciousness is raised, and the tension between our implicitly controlled system and the evidence of overgeneralization to which we have been made aware serves as the interface allowing system change" (N. C. Ellis, 2008, p. 240). Following this rationale, the next logical step would be to extend the investigation to awareness as 'reconstruction' and thus to determine "whether learners are aware of the knowledge they have stored and are retrieving" (R. Ellis, 2015, p. 421). Several methodological aspects would have to be considered with respect to the measurement of MLA within this new research endeavor: As previously mentioned, complementary measures such as confidence ratings or source attributions would counterbalance the verbal bias of the TAPs and of the THAM. With further reference to the valuable considerations of Rebuschat and his colleagues (2015), a subsequent step towards data triangulation would be to address the issue of 'reconstruction' by means of retrospective probing (see also Andringa & Rebuschat, 2015, p. 190). More precisely, post-task questionnaires or semi-directed interviews could shed light on the learners' capacity of retrieving knowledge about crosslinguistic correspondences from long-term memory.

Finally, apart from the specific approach(es) to measuring MLA and its beneficial effects

on learning processes and products, another consideration seems crucial for future investigations into the role of MLA in multilingual contexts: If we know that more analytic, explicit kinds of reasoning can help learners of multiple languages to make appropriate inferences based on the various linguistic resources that are available to them, then how do we provide them with the opportunities of developing MLA? This question leads directly to the implications that our findings might entail for teaching practices in multilingual classrooms. There is some evidence from classroom-based research that very strict applications of communicative language teaching, such as, for example, French immersion programs in Canada (e.g., Swain, 1985), have failed to provide learners with sufficient explicit knowledge to reconstruct certain grammatical features for subsequent output. Based on these observations, the exclusive focus on meaning in communicative classrooms was questioned by a number of researchers in instructed SLA, making way to the focus on form approach (Long, 1991; see also Nunan, 1998). Even if the term has been used to cover various kinds of more or less explicit classroom practices that have yielded somewhat different results in terms of promoting learner accuracy (Williams, 1995), SLA researchers agree on the usefulness of grammar instruction (R. Ellis, 2006, pp. 85–86). As previously outlined, this position is related to the limited success of a purely naturalistic approach to the acquisition of non-native languages. For example, Klein (1998) reported on a longitudinal investigation documented earlier (Perdue, 1993), where adult learners from different linguistic backgrounds acquired a grammatically simplified version of the target language from social immersion into the target language community. This so-called "Basic Variety" was characterized by an increasing number of content words, but a persisting absence of inflectional morphology (Klein, 1998, p. 544). Whereas Klein did not dismiss this Basic Variety as a deficient form of language use, his observations are mirrored by current assumptions on exclusively implicit learning conditions: "Although L2 learners are sourrounded by language, not all of it 'goes in'", which may be explained by the fact that "implicit tallying does not take place for low salient cues for which pattern recognition units have never been consolidated" (N. C. Ellis, 2015, p. 12). Moreover, in line with Schmidt's (1990) noticing hypothesis, communicative language teaching has been shown to benefit from a certain degree of form-focused instruction. Such more explicit approaches "aim at directing learners' attention to the problem area involved and thus assume that

providing learners with explicit knowledge about a specific linguistic feature can facilitate its eventual acquisition by aiding the process of noticing the form in the input" (N. C. Ellis & Cadierno, 2009, p. 125). Learners' attention can also be drawn to elements from another language in order to promote crosslinguistic awareness (see White & Horst, 2012).

If these considerations are extended to multilingual education, that is, instructional settings where learners have at least one other non-native language in their repertoire, different aspects ought to be taken into consideration. Namely, one of the goals to pursue should be to "relate the way multilinguals learn second/foreign languages to the way they communicate in real life, making it possible for the learner to activate the use of metalinguistic awareness and communicative competence acquired in previously learned languages in order to learn more efficiently" (Ruiz de Zarobe & Ruiz de Zarobe, 2015, p. 7). Moreover, future investigations into the nature and the effects of metalinguistic processes would also benefit from examining in more realistic learning situations (R. Ellis, 2012, p. 174). To what extent do learners consciously manipulate their different languages in classroom-based activities? Are crosslinguistic meaning-making strategies related to success? And what forms can they take in the communicative classroom? Different kinds of metalinguistic activities subsumed under the label of "languaging" have been documented in second language classrooms in Canada (Swain & Watanabe, 2013; Swain, 2006a). With the goal of mobilizing potentially profitable language learning experiences from a larger variety of languages, the concept has been extended to multilingual contexts under the label of "translanguaging" which typically refers to the "combination of two or more languages in a systematic way within the same learning activity" (Cenoz & Gorter, 2011, p. 359). Recent research in applied linguistics has mostly adopted a socially-informed research agenda to provide ample evidence for the positive implications that "translanguaging" activites might have in the construction of linguistic identity (e.g., Blackledge & Creese, 2014; García & Wei, 2014). However, it appears to be of crucial interest to examine such consciousness-raising activities in multilingual classrooms from a critical rationalist perspective. Namely, the ways in which multilinguals manipulate their languages in instructional settings can be diverse (e.g., Cenoz & Gorter, 2015). Such conscious manipulations appear to provide opportunities for learners to mobilize various kinds of

implicit and explicit background knowledge. In turn, based on the findings of the present study, future inquiries might focus on the different kinds of metalinguistic processes that emerge from specific classroom activities, and then investigate what kinds of explicit (or implicit) knowledge appear to be most conducive to further learning when a variety of background languages are available.

If we assume that explicit knowledge "consists of facts about language which learners know they know and can tell you they know" (R. Ellis, 2015, p. 419), then it must also be assumed that this knowledge can be more or less analyzed. For example, previous research with a focus on the use of metalanguage during a collaborative text reconstruction task has differentiated between more or less explicit kinds of metalinguistic processing (Fortune, 2005). In this particular study, metalanguage that was devoid of technical terminology was considered the most basic type. In this kind of collaborative dialogue, this would simply entail negociation of correctness or meaning, yet in absence of grammatical terms and without justifications based on underlying rules or textual logic. Moreover, a further distinction was made between (A) technical terms, i.e., grammatical categories, and two types of non-technical terms, where (B) referred to those that are commonly used for generalizations, such as 'position' to discuss clause syntax, and (C) to the learner's perception of correctness, such as 'it *sounds* right', or 'is this the *right* way to *use* it' (Fortune, 2005, p. 26). The results of this study revealed that all kinds of metalinguistic manipulations were associated with 'sustained engagement with a form', thus increasing the chances of retention and reproduction. However, it was also observed that justifications on the basis of more explicit generalizations (e.g., rule formulations) led to more efficient negociations of correctness or meaning. On the whole, this is what may be retained from these observations:

> The ultimate goal of explicit learning is not simply knowledge of concepts or rules. Instruction aiming for explicit learning is mostly provided to help learners process the input in a way that is conducive to the L2 acquisition process. (Andringa & Rebuschat, 2015, p. 189)

More specifically with respect to multilingual contexts, this would entail exploring the ways by which the learning process of a third or additional language may be complemented

by knowledge construction and reconstruction at the crossroads of their different languages. Not surprisingly, previous research on multilingual classroom activites has provided evidence for the multidirectionality of transfer phenomena, that is, transfer may take place between all of the learners' languages and in all directions (Cenoz & Gorter, 2011, p. 363). While providing fertile grounds for our research agenda, the variability with respect to each individual's learning experiences also raises further issues. As previously stated, the ability to verbalize is generally taken to reflect the process of becoming aware. Moreover, the use of literacy-based strategies in both reading and writing tasks was found to transcend linguistic borders (e.g., Cenoz & Gorter, 2011; Durgunoğlu, 2002; Schwartz et al., 2007). If we consider that the individuals in a multilingual classroom often come from various cultural, social and educational backgrounds and that their different languages may have been acquired in different surroundings, a number of subsequent questions emerge. For example, learners who do not read and write in their background languages are not expected to engage in metalinguistic processes in a way that would be comparable to bi- or multiliterate learners. In turn, if a longer-term goal was to make predictions about the benefits of particular types of form-focused instruction based on the metalinguistic processes observed in multilinguals, one would have to examine the role of the educational context in which previous languages were acquired and used. In a similar vein, the results of a recent study in the field of TLA, where parental education predicted the performance of a group of multilinguals on a written exam in L3, were interpreted as follows:

> From the evidence gathered it seems that parental education is an external factor which may influence the amount of metalinguistic knowledge and metalinguistic awareness that multilinguals may be able to gain, or at least parental education may have a more important role than was previously thought due to its known association with literacy development. (De Angelis, 2015, p. 14)

In order to tackle these complex questions, researchers might consider Hulstijn's (2015b) distinction between 'basic' and 'higher' language cognition to establish a sound theoretical basis concerning the contribution of literacy practices to specific aspects of proficiency in both native and non-native languages. Finally, we believe that the inherent complexity of multilingualism both at the individual and at the societal level calls for "partnerships of paradigms" (Mackey, 2014, p. 364). As indicated throughout the present discussion, future

investigations shall be devoted to interdisciplinarity:

> Although social psychology and applied linguistics might tackle multilingualism from different epistemological perspectives, from a methodological standpoint the study of multilingualism would benefit greatly from an interdisciplinary approach. (Comanaru & Dewaele, 2015, p. 12)

7   Conclusion

This last chapter summarizes the main aspects of the present research. The first subsection recapitulates the underlying goal of the study (7.1), followed by a review of the research protocol and its main findings (7.2). The third subsection reports on the major implications of the study (7.3), and the last one to its contribution to the fields of second and third language acquisition (7.4).

7.1   The underlying goal

The present study has examined the roles of metalinguistic awareness and of L2 proficiency in positive lexical transfer from English (L2) to German (L3) in the Francophone province of Québec. Moving away from a long-standing research tradition in which "interlanguage" was considered a deficient system that was characterized by interferences from the learner's native language (e.g., Corder, 1967; Nemser, 1969; Selinker, 1972), the focus of our study was on the positive influence of a non-native background language on the acquisition of a third or additional language. More precisely, we were interested in the conditions that might lead initial stage learners of German (L3) to make use of related background vocabulary in English (L2) to understand new words in German. In line with a growing body of research on individual multilingualism (e.g., Aronin & Singleton, 2012; De Angelis, 2007; Hammarberg, 2010; Hufeisen, 2000; Jessner, 2006), the acquisition of third or additional languages was considered as qualitatively different from that of second languages especially with respect to individual learning experiences which may be expected to enhance metalinguistic awareness (MLA). Furthermore, since heightened levels of MLA were shown to facilitate further language learning (e.g., Abu-Rabia & Sanitsky, 2010; Bono & Stratilaki, 2009; Jessner, 1999; Moore, 2006), the specific goal of the present study was to examine whether MLA might be stronger predictor of positive lexical transfer than two other variables that had been identified as crucial for transfer processes, namely proficiency and exposure levels in the source language (e.g., Dewaele, 2001; Odlin & Jarvis, 2004; Ringbom, 2007; Williams & Hammarberg, 1998).

7.2   The study in a nutshell

Positive lexical transfer was operationalized by the correct translations of unknown lexical items in German (L3) that could be traced back to an English cognate. Levels of MLA were established by means of a shortened version of the THAM (Test d'habiletés métalinguistiques) (Pinto & El Euch, 2015) and complemented by the qualitative analysis (see e.g., Leow, 1997) of think-aloud protocols (TAPs) that our participants were asked to produce throughout the translation task. Given the respective objects of inquiry and the types of metalinguistic processing required by each of the tasks, the results of the THAM were taken to represent a reflexive dimension of MLA, whereas as the TAPs reflected an applied dimension.[64] The Michigan Test of English Language Proficiency (MTELP) (Corrigan et al., 1979) was used to measure their proficiency in English (L2), while levels of L2 exposure were addressed in a background questionnaire (see Grosjean et al., 2000; Marian et al., 2007), which was further complemented by two questions aimed at establishing levels of interest in the German language and culture.

Two independent models of binomial logistic regressions, one of which featured the THAM as the exclusive measure of MLA while the other also included the MLA score from the TAPs, each identified MLA as the strongest predictor of positive lexical transfer from English, followed by English language proficiency. The *reflexive* dimension of MLA alone, as measured by the THAM, yielded a Chi-square statistic of 8.83 and a *p*-value of .003, which was followed closely by the MTELP ($X^2 = 6.63$, $p = .01$) in the first model. However, when the MLA score from the TAPs was entered into equation, this *applied* dimension of MLA was by far the strongest predictor of positive transfer, as evidenced by a particularly high statistic in this second model ($X^2 = 26.44$ $p < .001$), followed by the THAM ($X^2 = 4.55$, $p = .03$) whereas the MTELP did not reach the level of significance ($X^2 = 3.12$, $p = .08$). On the whole, our results point to the crucial role of MLA in the transfer process under investigation. Moreover, it appears that further investigations adopting a fine-grained methodological approach to the applied dimension of this complex construct might amplify its role in relation to other influential factors that may be expected to affect

---

[64] Refer to sections 4.5.1.2, 5.2.3, 6.1.2 and 6.2 for a detailed account of this conceptual distinction.

transfer processes in multilingual contexts. What may be retained from these findings will be briefly summarized for future investigations into language transfer, on the one hand, and for the role of explicit knowledge and learning in multilingual settings, on the other.

7.3   The major implications

For the study of "crosslinguistic interaction" – an umbrella term for all kinds of transfer phenomena (e.g., Bono, 2011; Jessner, 2008a) – the implications of our findings may be summed up as follows. Essentially, a more detailed account of cognitive variables related to the individual's learning experience might add substantially to the psycholinguistic investigation of transfer phenomena. Even if MLA has often been mentioned in studies on positive transfer (e.g., Gibson & Hufeisen, 2003; Peyer et al., 2010), it has, to our knowledge, never been measured as an explanatory variable along with other influential factors such as typology, proficiency, recency, etc. Especially in multilingual contexts where learning experiences and opportunities for crosslinguistic interaction are multiplied, Hufeisen's (2000) Factor Model, which was discussed in the conceptual framework of our study (see section 2.1.2.1, Figure 1) is thought to provide a sound basis for a conclusive analysis of the variables that may be expected to affect learning processes of all types. While the present study focused on the cognitive dimension of the transfer process, future research into crosslinguistic interaction might also want to provide a more comprehensive account of the affective dimension, an aspect that the present study did not seek to explore in depth (see section 6.1.5). However, as suggested by Hufeisen's (2000) Factor Model, affective variables can be expected to impact the interaction of multiple languages in the acquisition process.

More generally, with respect to the role of explicit knowledge and learning in language development, our findings are consistent with a large body of research from different fields having established a prominent role of MLA in the acquisition process of first, second and third languages (e.g., Bialystok, 2001a; El Euch, 2010; Elder & Manwaring, 2004; Gombert, 1990; Jessner, 2006; Karmiloff-Smith, 1992; Rauch et al., 2012; Roehr, 2007). More specifically, our results endorse previous research findings in identifying MLA as one of the major constituents of multilingual development (e.g., Cenoz, 2003a; Cook, 2007;

Herdina & Jessner, 2002; Jessner, 2008a). However, from a methodological perspective, the notion of MLA has been treated in a variety of ways, which makes it difficult to generalize about the concept itself and the impact it is thought to have on different learning processes. The crucial point to be retained from the present study is that a rigorous measurement of MLA must be grounded in a thorough analysis of its theoretical and conceptual underpinnings.

For example, certain researchers adopting a holistic approach to multilingual development have used the term MLA to describe a variety of mental processes involving the learners' explicit manipulation of their own languages in use, yet without operationalizing the construct for empirical inquiry (e.g., Jessner, 1999, 2005; Moore, 2006; Volgger, 2010). Out of the studies designed for a systematic measurement of MLA in applied linguistics, many have focused on the mastery of metalinguistic knowledge, mostly operationalized as the correction and description of errors (e.g., Elder & Manwaring, 2004; Roehr & Gánem-Gutiérrez, 2009; Zietek & Roehr, 2011) or strictly as knowledge of appropriate metalinguistic terms, including names for grammatical categories (e.g., Alderson et al., 1997). Others have aimed at a more comprehensive account of metalinguistic abilities, by focusing on the manipulation (and justification) of a larger variety of linguistic features, including semantic and pragmatic aspects of language use (e.g., Pinto & El Euch, 2015; Pinto, Melongo, & Iliceto, 2011; Pinto et al., 1999). However, in the majority of these studies, the notion of awareness has been treated in terms of an ability that becomes manifest (explicit) in the learners' (verbal) accounts of 'what they know about specific structural aspects of the language'. Moments in which learners fail to report on certain structures in a certain way are thus interpreted in terms of a lack of knowledge. To address this conceptual and methodological issue, Rebuschat and his colleagues (2015) have suggested various measures of awareness, whose triangulation may be expected to provide a more complete account of the concept of MLA.

7.4   The contribution of the present study

In the present study, the concept of *awareness* was understood in terms of a mental activity that involves the temporary focus of attention (e.g., Bialystok, 2001). What is implied in the

part of the concept labeled *metalinguistic* is the deliberate control of such attention over parts of the language itself (see Gombert, 1990). In this sense, the notion of MLA, as we conceptualized it, does not (necessarily) correspond to depth of analysis – in terms of levels of abstraction within the mental process, i.e., what kind of underlying knowledge is accessible to be consciously manipulated – nor does it (necessarily) imply a specific degree of explicitness, i.e., the extent to which knowledge is verbalized and what terms are used to describe it. Given this broad definition of metalinguistic awareness, initial stages of awareness were included in the operationalization of the variable. Moreover, we adopted a mixed approach to its measurement, in order to highlight two distinct contexts of language use, in which metalinguistic awareness served two different purposes. On the one hand, we adopted the THAM-3 (Pinto & El Euch, 2015) to measure the participants' overall capacity to describe different features of their native language at varying degrees of analysis and explicitness. This is what we called the *reflexive* dimension of MLA, as opposed to the *applied* awareness that was examined in the TAPs. The protocol that was adopted to investigate the metalinguistic processes that our participants engaged in when faced with unknown vocabulary in their target language German (L3) was largely inspired by research in the field of cognitive psychology (e.g., Hama & Leow, 2010; Leow et al., 2011; Rebuschat et al., 2015; Rosa & Leow, 2004; Sachs & Suh, 2007). In line with Leow's (2000) conception of "low" awareness at the level of *noticing* and "high" awareness at the level of *understanding* (see also Schmidt, 1990, 2010), the TAPs were analyzed for levels of awareness ranging from an initial stage of "intuitive" awareness to the most elaborate stage, where learners manipulated their linguistic resources to a degree of abstraction enabling them to make generalizations based on underlying rules. In focusing on the active application of crosslinguistic meaning-making strategies, we concentrated on explicit learning, rather than on a certain type of explicit (metalinguistic) knowledge that results from the learning process. Having thus combined different methods to shed light on metalinguistic processes, the results of the present study endorse previous observations regarding the role of explicit learning in instructional settings:

> Many learners – especially adult learners – possess substantial metalinguistic knowledge of the L2. While there is uncertainty regarding the extent to which such knowledge is important for developing L2 proficiency, there is clearer evidence that

it assists learners to focus on form when performing tasks. (R. Ellis, 2012, p. 175)

In the specific context of our study, the above observation may be extended to all of the learner's languages. Our results revealed that higher levels of applied awareness were conducive to the process of making meaning of new vocabulary in L3. In other words, depth of analysis and degree of explicitness were in fact associated with more appropriate form-meaning mappings, whereas the task itself did not necessitate more than an intuitive awareness of crosslinguistic correspondences to lead our participants to resort to positive transfer from English. Even if the reflexive and the applied dimensions of MLA are thought to bear a reciprocal relationship (see sections 5.2.3, 6.1.2 and 6.2), our findings suggest that it is the process of becoming aware of certain crosslinguistic correspondences and the extent to which existing metalinguistic knowledge is integrated into this process that deserves further attention in the study of multiple language acquisition.

> Learners' language systematically emerges from their history of interactions of implicit and explicit language learning, from the statistical abstraction of patterns latent within and across forms and functions in language usage. (N. C. Ellis, 2015, p. 14)

In order to gain further insight into these interactions when various languages are at play, future research into multilingual development might ultimately benefit from what Ortega calls "transdisciplinarity", that is "an attempt to solve complex knowledge problems with social impact by working across multiple disciplinary boundaries as well as with social actors outside academia" (Ortega, 2013, p. 3).

REFERENCES

Abu-Rabia, S., & Sanitsky, E. (2010). Advantages of bilinguals over monolinguals in learning a third language. *Bilingual Research Journal*, *33*(2), 173–199.

Alderson, J. C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, *1*, 93–121.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition*, *37*, 185–196.

Ard, J., & Homburg, T. (1983). Verification of language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning. Issues in second language research* (pp. 157–176). Rowley, MA: Newbury House.

Aronin, L., & Hufeisen, B. (Eds.). (2009). *The exploration of multilingualism: Development of research in L3, multilingualism and multiple language acquisition*. Amsterdam/Philadelphia: Benjamins.

Aronin, L., & Singleton, D. (2008). Multilingualism as a new linguistic dispensation. *International Journal of Multilingualism*, *5*(1), 1–16.

Aronin, L., & Singleton, D. (2012). *Multilingualism*. Amsterdam/Philadelphia: Benjamins.

Banta, F. (1981). Teaching German vocabulary: The use of English cognates and loan words. *Modern Language Journal*, *65*(2), 129–136.

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, *23*(4), 459–484.

Bayona, P. (2009). *Crosslinguistic influences in the acquisition of Spanish L3*. Unpublished doctoral dissertation. University of Western Ontario. London, ON.

Bernardini, S. (2001). Think-aloud protocols in translation research. *Target*, *13*(2), 241–263.

Bialystok, E. (1993). Metalinguistic awareness: The development of children's representations of language. In C. Pratt & A. F. Garton (Eds.), *Systems of representation in children: Development and use* (pp. 211–233). Chichester, UK: Wiley.

Bialystok, E. (2001a). *Bilingualism in development: Language, literacy and cognition*. New York: Cambridge University Press.

Bialystok, E. (2001b). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, *21*, 169–181.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11.

Bialystok, E., Craik, F., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290–303.

Bialystok, E., Peets, K. F., & Moreno, S. (2014). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied Psycholinguistics*, *35*, 177–191.

Bild, E.-R., & Swain, M. (1989). Minority language students in a French Immersion programme: Their French proficiency. *Journal of Multilingual and Multicultural Development*, *10*(3), 255–274.

Blackledge, A., & Creese, A. (Eds.). (2014). *Heteroglossia as practice and pedagogy*. New York: Springer.

Bono, M. (2011). Crosslinguistic interaction and metalinguistic awareness in third language acquisition. In G. De Angelis & J.-M. Dewaele (Eds.), *New trends in crosslinguistic influence and multilingualism research* (pp. 25–52). Bristol: Multilingual Matters.

Bono, M., & Stratilaki, S. (2009). The M-factor, a bilingual asset for plurilinguals? Learners' representations, discourse strategies and third language acquisition in institutional contexts. *International Journal of Multilingualism*, *6*(2), 207–228.

Bowles, M. A. (2008). Task type ands reactivity of verbal reports in SLA: A first look at a L2 task other than reading. *Studies in Second Language Acquisition*, *30*(3), 359–387.

Bowles, M. A. (2010). *The think-aloud controversy in second-language research*. New York: Routledge.

Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, *27*, 415–440.

Burton, G. (2013). Cross-linguistic influence in non-native languages: Explaining lexical transfer using language production models. *International Journal of Multilingualism*, *10*(1), 46–59.

Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216–236). New York: Routledge.

Caspari, D., & Rössler, A. (2008). Französisch gegen Spanisch? Überlegungen aus Sicht der romanischen Mehrsprachigkeitsdidaktik. *Zeitschrift für Fremdsprachenforschung*, *19*(1), 61–82.

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 8–20). Clevedon: Multilingual Matters.

Cenoz, J. (2003a). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, *7*(1), 71–87.

Cenoz, J. (2003b). The role of typology in the organization of the multilingual lexicon. In J. Cenoz, U. Jessner & B. Hufeisen (Eds.), *The multilingual lexicon* (pp. 103–116). Dordrecht: Kluwer.

Cenoz, J., & Gorter, D. (2011). Focus on multilingualism: A study of trilingual writing. *The*

Modern Language Journal, *95*, 356–369.

Cenoz, J., & Gorter, D. (2015). Towards a holistic approach in the study of multilingual education. In J. Cenoz & D. Gorter (Eds.), *Multilingual education: Between language learning and translanguaging* (pp. 1–15). Cambridge: Cambridge University Press.

Cenoz, J., Hufeisen, B., & Jessner, U. (Eds.). (2001). *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Clevedon: Multilingual Matters.

Cenoz, J., Jessner, U., & Hufeisen, B. (Eds.). (2003). *The multilingual lexicon*. Dordrecht: Kluwer.

Cenoz, J., & Valencia, J. F. (1994). Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics*, *15*(2), 195–207.

Chomsky, N. (1976). *Reflections on language*. London: Temple Smith.

Clark, E. V. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. Jarvella & W. J. M. Levelt (Eds.), *The child's conception of language.* (pp. 17–43). New York: Springer.

Clark, E. V., & Andersen, E. S. (1979). Spontaneous repairs: Awareness in the process of acquiring language (Paper presented at the Biennial Meeting of the Society for Research in Child Development, San Francisco, March 1979). *Papers & Reports on Child Language Development*, *16*, 1–12.

Comanaru, R.-S., & Dewaele, J.-M. (2015). A bright future for interdisciplinary multilingualism research. *International Journal of Multilingualism*, *12*, 1–15.

Cook, V. (1992). Evidence for multicompetence. *Language Learning*, *42*, 557–591.

Cook, V. (2002). *Portraits of the L2 user*. Clevedon: Multilingual Matters.

Cook, V. (2007). Multi-competence: Black hole or wormhole for second language acquisition research? In Z.-H. Han (Ed.), *Understanding second language process* (pp. 16–26). Clevedon: Multilingual Matters.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, *5*(4), 161–170.

Corder, S. P. (1979). Language distance and the magnitude of the language learning task. *Studies in Second Language Acquisition*, *2*, 27–36.

Corrigan, A., Dobson, B., Kellman, E., Spaan, M., Strowe, L., & Tyma, S. (1979). *Michigan Test of English Language Proficiency*. Ann Arbor: University of Michigan.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, *9*(8), 137–151.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Coupland, N., & Jaworski, A. (2004). Sociolinguistic perspectives on metalanguage: Reflexivity, evaluation and ideology. In A. Jaworski, N. Coupland & D. Galasinski (Eds.), *Metalanguage: Social and ideological perspectives* (pp. 15–52). Berlin:

Mouton De Gruyter.

Culioli, A. (1968). La formalisation en linguistique. *Cahiers pour l'analyse*, *9*, 106–117.

Cummins, J. (1991). Conversational and academic language proficiency in bilingual contexts. In J. H. Hulstijn & J. F. Matter (Eds.), *Reading in two languages* (pp. 75–89). Amsterdam: AILA.

Cummins, J. (2005). Teaching for cross-language transfer in dual language education: Possibilities and pitfalls. *TESOL Symposium on Dual Language Education: Teaching and Learning Two Languages in the EFL Setting*, 1–18.

De Angelis, G. (2005). Multilingualism and non-native lexical transfer: An identification problem. *International Journal of Multilingualism*, *2*(1), 1–25.

De Angelis, G. (2007). *Third or additional language acquisition*. Clevedon: Multilingual Matters.

De Angelis, G. (2015). English L3 learning in a multilingual context: The role of parental education and L2 exposure within the living community. *International Journal of Multilingualism*, 1–18.

De Angelis, G., & Selinker, L. (2001). Interlanguage transfer and competing linguistic systems in the multilingual mind. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 42–58). Clevedon: Multilingual Matters.

de Bot, K. (1992). A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics*, *13*(1), 1–24.

de Bot, K. (2004). The multilingual lexicon: Modeling selection and control. *International Journal of Multilingualism*, *1*, 1–16.

de Bot, K., & Lowie, W. (2010). On the stability of representations in the multilingual lexicon. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind, Converging evidence in language and communication research (CELCR)* (pp. 117–134). Amsterdam: Benjamins.

de Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56.

DeKeyser, R. (2014). The philosphy of science and the social-cognitive dichotomy in research in language learning and teaching. *Studies in Second Language Acquisition*, *36*(3), 365–368.

Dewaele, J.-M. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Applied Linguistics*, *19*(4), 471–490.

Dewaele, J.-M. (2001). Activation or inhibition? The interaction of L1, L2 and L3 on the language mode continuum. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 69–89). Clevedon: Multilingual Matters.

Dillon, A. M. (2009). Metalinguistic awareness and evidence of cross-linguistic influence among bilingual learners in Irish primary schools. *Language Awareness*, *18*(2), 182–197.

Dörnyei, Z. (1997). Psychological processes in cooperative language learning: Group dynamics and motivation. *The Modern Language Journal*, *81*(4), 482–493.

Dressler, C., Carlo, M. S., Snow, C. E., August, D., & White, C. E. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14*(2), 243–255.

Durgunoğlu, A. Y. (1997). Bilingual reading: Its components, development, and other issues. In A. M. B. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 255–276). New Jersey: Erlbaum.

Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, *52*(1), 189–204.

Ecke, P. (2008). Cross-linguistic influence on word search in tip-of-the-tongue states. *TESOL Quarterly*, *42*(3), 515–527.

Ecke, P. (2015). Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language and Cognition*, *18*(2), 145–162.

El Euch, S. (2010). Attitudes, motivations et conscience métalinguistique chez des bilingues et des trilingues adultes: effets, similarités et différences. *Language Awareness*, *19*(1), 17–33.

El Euch, S. (2011). De la typologie de la bilingualité à une typologie du plurilinguisme ou de la multilingualité : un hommage à Josiane Hamers. *Canadian Modern Language Review*, *67*(1), 55–90.

Elder, C., & Manwaring, D. (2004). The relationship between metalinguistic knowledge and learning outcomes among undergraduate students of Chinese. *Language Awareness*, *13*(3), 145–162.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*(1), 1–24.

Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal*, *92*(2), 232–249.

Ellis, N. C. (2015). Implicit AND explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). Amsterdam/Philadelphia: Benjamins.

Ellis, N. C., & Cadierno, T. (2009). Constructing a second language. *Annual Review of Cognitive Linguistics*, *7*(Special section), 111–290.

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2015). Usage-based language learning. In B. MacWhinney & W. O'Grady (Eds.), *Handbook of language emergence* (pp. 163–180). Malden, MA: Wiley.

Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics*, *6*, 118–131.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *Teaching Notes on Population*, *40*(1), 83–107.

Ellis, R. (2012). Focus on the learner. In *Language teaching research and language pedagogy* (pp. 151–193). Hoboken: Wiley-Blackwell.

Ellis, R. (2015). Form-focused instruction and the measurement of implicit and explicit L2 knowledge. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 417–441). Amsterdam/Philadelphia: Benjamins.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

European Commission. (2008). *Multilingualism: An asset for Europe and a shared commitment*.

Faerch, C., & Kasper, G. (Eds.). (1987). *Introspection in second language research*. Clevedon: Multilingual Matters.

Falk, Y., & Bardel, C. (2010). The study of the role of the background languages in third language acquisition. The state of the art. *IRAL - International Review of Applied Linguistics in Language Teaching*, *48*, 185–219.

Fehling, S. (2008). *Language Awareness und bilingualer Unterricht: Eine komparative Studie [Language awareness and bilingual instruction: A comparative study]* (2nd ed.). Frankfurt/Main, Germany: Peter Lang.

Filatova, K. (2010). Third language acquisition, macrocategories and synonymy. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind. Converging evidence in language and communication research (CELCR)* (pp. 85–96). Amsterdam: Benjamins.

Fitzpatrick, T. (2006). Habits and rabbits. Word associations and the L2 lexicon. *EUROSLA Yearbook*, *6*, 121–145.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441.

Forsyth, H. (2014). The influence of L2 transfer on L3 English written production in a bilingual German/Italian population: A study of syntactic errors. *Open Journal of Modern Linguistics*, *4*, 429–456.

Fortune, A. (2005). Learners' use of metalanguage in collaborative form-focused L2 output tasks. *Language Awareness*, *14*(1), 21–38.

Franceschini, R. (2009). The genesis and development of research in multilingualism: Perspectives for future research. In L. Aronin & B. Hufeisen (Eds.), *The exploration of multilingualism: Development of research in L3, multilingualism and multiple language acquisition* (pp. 28–61). Amsterdam/Philadelphia: Benjamins.

Frawley, W., & Lantolf, J. P. (1985). Second language discourse : A Vygotskyan perspective. *Applied Linguistics*, *6*(1), 19–44.

Friel, B. M., & Kennison, S. M. (2001). Identifying German–English cognates, false cognates, and non-cognates: Methodological issues and descriptive norms. *Bilingualism: Language and Cognition*, *4*(3), 249–274.

Gabryś-Barker, D. (2006). Language activation in the thinking processes of a multilingual language user. *International Journal of Multilingualism*, *3*(2), 105–124.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Basingstoke: Palgrave Macmillan.

García-Mayo, M. del P. (2012). Cognitive approaches to L3 acquisition. *International Journal of English Studies*, *12*(1), 129–146.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.

Gibson, M., & Hufeisen, B. (2003). Investigating the role of prior foreign language knowledge. In J. Cenoz, U. Jessner & B. Hufeisen (Eds.), *The multilingual lexicon* (pp. 87–102). Dordrecht: Kluwer.

Gibson, M., & Hufeisen, B. (2006). Metalinguistic processing control mechanisms in multilingual learners of English. *International Journal of Multilingualism*, *3*(2), 139–153.

Gombert, J. E. (1990). *Le développement métalinguistique*. Paris: Presses Universitaires de France.

Gombert, J. E. (2000). Développement métalinguistique, lecture et illettrisme. Retrieved November 11, 2013, from http://adapt-scol-franco.educ.infinit.net

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210–223.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. M. B. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Mahwah, NJ: Erlbaum.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Oxford: Blackwell.

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & L. Ping (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). Malden, MA: Blackwell.

Grosjean, F. (2014). Bicultural bilinguals. *International Journal of Bilingualism*, Published online, 1 April 2014. DOI: 10.1177/1367006914526297.

Grosjean, F., Jaccard, R., & Cividin, V. (2000). *Questionnaire pour personnes bilingues*. Université de Neuchâtel.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, *32*, 465–491.

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 21–41). Clevedon: Multilingual Matters.

Hammarberg, B. (2010). The languages of the multilingual: Some conceptual and terminological issues. *IRAL - International Review of Applied Linguistics in Language Teaching*, *48*, 91–104.

Hardin, J., & Hilbe, J. (2007). *Generalized linear models and extensions* (2nd ed.). College Station: Stata Press.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Clevedon: Multilingual Matters.

Hondo, J. (2013). Synthesizing social and cognitive approaches in SLA. *Working Papers in Educational Linguistics*, *29*(2), 1–23.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley.

Hu, G. (2011). Metalinguistic knowledge, metalanguage, and their relationship in L2 learners. *System*, *39*, 63–77.

Hufeisen, B. (2000). A European perspective: Tertiary languages with a focus on German as L3. In J. Rosenthal (Ed.), *Handbook of undergraduate second language education: English as a second language, bilingual, and foreign language instruction for a multilingual world* (pp. 209–229). Mahwah, NJ: Erlbaum.

Hufeisen, B., & Gibson, M. (2003). Zur Interdependenz emotionaler und kognitiver Faktoren im Rahmen eines Modells zur Beschreibung sukzessiven multiplen Sprachenlernens. *Bulletin Suisse de Linguistique Appliquée*, *78*, 13–33.

Hufeisen, B., & Marx, N. (2007). *EuroComGerm - Die sieben Siebe: Germanische Sprachen lesen lernen*. Aachen: Shaker.

Hulstijn, J. H. (2014). Epistemological remarks on a social-cognitive gap in the study of second language learning and teaching. *Studies in Second Language Acquisition*, *36*(3), 375–382.

Hulstijn, J. H. (2015a). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning: The virtues and pitfalls of a two-system view. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 25–46). Amsterdam/Philadelphia: Benjamins.

Hulstijn, J. H. (2015b). *Language proficiency in native and non-native speakers: Theory and research*. Amsterdam: Benjamins.

Jakobson, R. (1960). Closing statements: Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, *50*(2), 245–309.

Jarvis, S. (2003). Probing the effects of the L2 on the L1: A case study. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 81–102). Clevedon: Multilingual Matters.

Jarvis, S. (2009). Lexical transfer. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 99–124). Clevedon: Multilingual Matters.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Jessner, U. (1999). Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language Awareness*, *8*(3-4), 201–209.

Jessner, U. (2005). Multilingual metalanguage, or the way multilinguals talk about their languages. *Language Awareness*, *14*(1), 56–68.

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh: Edinburgh University Press.

Jessner, U. (2008a). A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal*, *92*, 270–283.

Jessner, U. (2008b). Language awareness in multilinguals: Theoretical trends. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 6, pp. 1–13). New York: Springer.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from metalinguistic and repair data. *Cognition*, *23*, 95–147.

Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kellerman, E. (1977). Towards a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin*, *2*, 58–145.

Kellerman, E. (1983). Now you see it. Now you don't. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112–134). Newbury: Rowley.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics*, *15*, 125–150.

Klein, W. (1998). The contribution of second language acquisition research. *Language Learning*, *48*(4), 527–550.

Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

Kuile, H., Veldhuis, M., van Veen, S. C., & Wicherts, J. M. (2011). Bilingual education, metalinguistic awareness, and the understanding of an unknown language. *Bilingualism: Language and Cognition*, *14*(2), 233–242.

Lado, R. (1957). *Linguistics across cultures. Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*(2), 141–165.

Larsen-Freeman, D. (2007). Reflecting on the cognitive-social debate in second language acquisition. *The Modern Language Journal*, *91*, 773–787.

Larsen-Freeman, D. (2012). Complexity theory. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 73–87). New York: Routledge.

Lasagabaster, D. (2001). The effect of knowledge about the L1 on foreign language skills and grammar. *International Journal of Bilingual Education and Bilingualism*, *4*(5), 310–331.

Lefrançois, P. (2000). *Capacités métalinguistiques en langues maternelle et seconde dans l'apprentissage de la lecture en français chez des hispanophones*. Unpublished doctoral dissertation. Université de Montréal. Montréal, QC.

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, *47*(3), 467–505.

Leow, R. P. (2000). A study of the role of awareness in foreign language behavior. *Studies in Second Language Acquisition*, *22*, 557–584.

Leow, R. P. (2015). Implicit learning in SLA: Of processes and products. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 47–67). Amsterdam/Philadelphia: Benjamins.

Leow, R. P., Johnson, E., & Zarate-Sandez, G. (2011). Getting a grip on the slippery construct of awareness: Toward a finer-grained methodological perspective. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 61–72). Georgetown University Press.

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, *26*(1), 35–57.

Lindqvist, C. (2009). The use of the L1 and the L2 in French L3: Examining cross-linguistic lexemes in multilingual learners' oral production. *International Journal of Multilingualism*, *6*(3), 281–297.

Lindqvist, C. (2010). Inter- and intralingual lexical influences in advanced learners' French L3 oral production. *IRAL - International Review of Applied Linguistics in Language Teaching*, *48*, 131–157.

Llama, R., Cardoso, W., & Collins, L. (2010). The influence of language distance and language status on the acquisition of L3 phonology. *International Journal of Multilingualism*, *7*(1), 39–57.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In

K. de Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspectives* (pp. 39–52). Amsterdam: Benjamins.

Mackey, A. (2014). Exploring questions of balance in interaction research. *Studies in Second Language Acquisition*, *36*(3), 380–383.

MacLennan, H. (1945). *Two solitudes*. Toronto, ON: Collins.

Malakoff, M. (1992). Translation ability: A natural bilingual and metalinguistic skill. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 515–530). Hillsdale, NJ: Erlbaum.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.

Marshall, J. C., & Morton, J. (1978). On the mechanics of EMMA. In A. Sinclair, R. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 225–239). New York: Springer.

Marx, N., & Mehlhorn, G. (2010). Pushing the positive: Encouraging phonological transfer from L2 to L3. *International Journal of Multilingualism*, *7*(1), 4–18.

Meara, P. (2006). Emergent properties of multilingual lexicons. *Applied Linguistics*, *27*(4), 620–644.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam/Philadelphia: Benjamins.

Meuter, R., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.

Milton, J. (2006). *Measuring second language vocabulary acquisition*. Bristol/Buffalo: Multilingual Matters.

Moore, D. (2006). Plurilingualism and strategic competence in context. *International Journal of Multilingualism*, *3*(2), 125–138.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nayak, N., Hansen, N., Krueger, N., & McLaughlin, B. (1990). Language-learning strategies in monolingual and multilingual adults. *Language Learning*, *40*(2), 221–244.

Nemser, W. (1969). *Approximative systems of foreign language learners. IRAL - International Review of Applied Linguistics in Language Teaching*, 9, 115-224.

Nilsson, A. (2009). La lecture des mots translinguistiques en français L3. *EUROSLA Yearbook*, *9*, 132–159.

Nunan, D. (1998). Teaching grammar in context. *ELT Journal*, *52*(7), 101–109.

Ó Laoire, M. (2006). Transfer and processing in multilingual learners. *International Journal of Multilingualism*, *3*(2), 86–87.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Odlin, T. (2003). Cross-linguistic influence. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 436–486). Malden, MA: Blackwell.

Odlin, T., & Jarvis, S. (2004). Same source, different outcomes: A study of Swedish influence on the acquisition of English in Finland. *International Journal of Multilingualism*, *1*(2), 123–140.

Ortega, L. (2011). SLA after the social turn: Where cognitivism and its alternatives stand. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 167–179). New York: Routledge.

Ortega, L. (2013). SLA for the 21st century: Disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Language Learning*, *63*, 1–24.

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual speakers: The Activation Threshold Hypothesis. *International Journal of Psycholinguistics*, *9*, 133–145.

Peal, E., & Lambert, W. (1962). The relation of bilingualism to intelligence. *Psychological Monographs*, *76*, 1–23.

Perdue, C. (1993). *Adult language acquisition: Cross-linguistic perspectives*. Cambridge: Cambridge University Press.

Peyer, E., Kaiser, I., & Berthele, R. (2010). The multilingual reader: Advantages in understanding and decoding German sentence structure when reading German as an L3. *International Journal of Multilingualism*, *7*(3), 225–239.

Piaget, J. (1929). *The child's conception of the world*. London: Routledge.

Pinto, M. A., & El Euch, S. (2015). *La conscience métalinguistique: Théorie, développement et instruments de mesure*. Québec, QC: Les Presses de l'Université Laval.

Pinto, M. A., Melongo, S., & Iliceto, P. (2011). Assessing metaphor comprehension as a metasemantic ability in students from 9-to-14 years-old. *LINGVARVM ARENA*, *2*, 57–77.

Pinto, M. A., Titone, R., & Trusso, F. (1999). *Metalinguistic awareness. Theory, development and measurement instruments*. Roma: Instituto Editoriali e Poligrafici Internazionali.

Pinto, M. A., Trusso, F., & Kristiansen, K. (2002). Metalinguistic abilities in Italian-English adult bilinguals: A comparison with Italian-speaking and English-speaking monolinguals. *Rivista Di Psicolinguistica Applicata*, *II*(2-3), 77–90.

Qasem, M., & Foote, R. (2010). Crosslanguage lexical activation. A test of the revised hierarchical and morphological decomposition models in Arabic-English bilinguals. *Studies in Second Language Acquisition*, *32*(01), 111–140.

Ransdell, S., Barbier, M.-L., & Niit, T. (2006). Metacognitions about language skill and

working memory among monolingual and bilingual college students: When does multilingualism matter? *International Journal of Bilingual Education and Bilingualism*, *9*(6), 728–741.

Ranta, L. (2002). The role of learners' language analytic ability in the foreign language communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159–180). Amsterdam: Benjamins.

Ranta, L., & Meckelborg, A. (2013). How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *Canadian Modern Language Review*, *69*(1), 1–33.

Rast, R. (2010). The use of prior linguistic knowledge in the early stages of L3 acquisition. *IRAL - International Review of Applied Linguistics in Language Teaching*, *48*, 159–183.

Rauch, D. P., Naumann, J., & Jude, N. (2012). Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish-German bilinguals. *International Journal of Bilingualism*, *16*(4), 402–418.

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, *24*, 146–161.

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of languages*. Amsterdam/Philadelphia: Benjamins.

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness. *Studies in Second Language Acquisition*, *37*, 299–334.

Remysen, W. (2004). Le recours au stéréotype dans le discours sur la langue française et l'identité québécoise : Une étude de cas dans la région de Québec [Appealing to stereotypes in discourse on the French language et Québécois identity: A case study in the Quebec City are. In D. Deshaies & D. Vincent (Eds.), *Discours et constructions identitaires [Discourse and identity constructions]* (pp. 95–121). Québec, QC: Les Presses de l'Université Laval.

Renou, J. (2001). An examination of the relationship between metalinguistic awareness and second-language proficiency of adult learners of French. *Language Awareness*, *10*, 248–267.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.

Ringbom, H. (1992). On L1 transfer in L2 comprehension and L2 production. *Language Learning*, *42*(1), 85–112.

Ringbom, H. (2001). Lexical transfer in L3 production. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 59–68). Clevedon: Multilingual Matters.

Ringbom, H. (2007). *The importance of cross-linguistic similarity in foreign language learning: Comprehension, learning and production*. Clevedon: Multilingual Matters.

Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign

language learning. In M. H. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 106–118). New York: Wiley-Blackwell.

Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge.

Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, *29*(2), 173–199.

Roehr, K., & Gánem-Gutiérrez, G. A. (2009). The status of metalinguistic knowledge in instructed adult L2 learning. *Language Awareness*, *18*(2), 165–181.

Rosa, E., & Leow, R. P. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics*, *25*, 269–292.

Rosa, E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, *21*, 511–556.

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, *27*(1), 107–127.

Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, *26*(2), 189–218.

Ruiz de Zarobe, L., & Ruiz de Zarobe, Y. (2015). New perspectives on multilingualism and L2 acquisition: An introduction. *International Journal of Multilingualism*, *12*(4), 1–11.

Ryan, E. B., & Ledger, J. W. (1984). Learning to attend to sentence structure: Links between metalinguistic development and reading. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 149–171). New York: Springer.

Sachs, R., & Suh, B.-R. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 197–227). Oxford: Oxford University Press.

Sanders, M., & Meijers, G. (1995). English as L3 in elementary school. *ITL Review of Applied Linguistics*, *107*(8), 59–78.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11–26.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, … I. Walker (Eds.), *Proceedings of CLaSIC 2010* (Vol. 4, pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J. H., Snellings, P., Simis, A., &

Stevenson, M. (2002). Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2 and EFL writing: A structural equation modelling approach. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 101–122). Dordrecht: Kluwer.

Schwartz, M., Geva, E., Share, D. L., & Leikin, M. (2007). Learning to read in English as third language: The cross-linguistic transfer of phonological processing skills. *Written Language & Literacy*, *10*(1), 25–52.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.

Selinker, L. (1969). Language transfer. *General Linguistics*, *9*, 67–92.

Selinker, L. (1972). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, *10*(3), 209–231.

Serrano, R. (2011). From metalinguistic instruction to metalinguistic knowledge, and from metalinguistic knowledge to performance in error correction and oral production tasks. *Language Awareness*, *20*(1), 1–16.

Singleton, D., & Aronin, L. (2007). Multiple language learning in the light of the theory of affordances. *Innovation in Language Learning and Teaching*, *1*(1), 83–96.

Singleton, D., Fishman, J., Aronin, L., & O Laoire, M. (Eds.). (2013). *Current multilingualism: A new linguistic dispensation*. New York: Mouton De Gruyter.

Singleton, D., & Ó Laoire, M. (2006). Psychotypologie et facteur L2 dans l'influence translexicale. *Acquisition et Interaction En Langue étrangère*, *24*, 1–12.

Smagorinsky, P. (1998). Thinking and speech and protocol analysis. *Mind, Culture and Activity*, *5*(3), 157–177.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford: Oxford University Press.

Swain, M. (2006a). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95–108). London, England: Continuum.

Swain, M. (2006b). Verbal protocols: What does it mean for research to use speaking as a data collection tool. In M. Chalhoub-Deville, C. A. Chapelle & P. A. Duff (Eds.), *Inference and generalizability in applied linguistics: Multiple perspectives* (pp. 97–113). Amsterdam/Philadelphia: Benjamins.

Swain, M., & Watanabe, Y. (2013). Languaging: Collaborative dialogue as a source of second language learning. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–8). Malden, MA: Wiley-Blackwell.

Tape, T. G. (n.d.). Interpreting diagnostic tests. Retrieved September 1, 2015, from http://gim.unmc.edu/dxtests/roc3.htm

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, *4*(2), 142–163.

Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.

Tremblay, M.-C. (2006). Cross-linguistic influence in third language acquisition: The role of L2 proficiency and L2 exposure. *Ottawa Papers in Linguistics*, *34*, 109–119.

Trudgill, P. (2000). *Sociolinguistics: An introduction to language and society* (4th ed.). London, England: Penguin.

Tunmer, W. E., & Herriman, M. (1984). The development of metalinguistic awareness: A conceptual overview. In W. E. Tunmer, C. Pratt & M. L. Herriman (Eds.), *Metalinguistic awareness in children: Theory, research and implications* (pp. 12–35). Berlin: Springer.

Tunmer, W. E., Pratt, C., & Herriman, M. (Eds.). (1984). *Metalinguistic awareness in children: Theory, research and implications*. Berlin: Springer.

Van Heuven, W., Dijkstra, T., & Grainger, J. (1998). Orthographic neigborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*, 468–483.

Volgger, M.-L. (2010). "Wenn man mehrere Sprachen kann, ist es leichter, eine weitere zu lernen..." Einblicke in die Mehrsprachigkeitsbewusstheit lebensweltlich mehrsprachiger Franzosen. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, *15*(2), 169–198.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Webb, S. (2007). The effects of synonymy on second-language vocabulary learning. *Reading in a Foreign Language*, *19*(2), 120–136.

Wei, L. (2006). The multilingual mental lexicon and lemma transfer in third language learning. *International Journal of Multilingualism*, *3*(2), 88–104.

Weinreich, U. (1953). *Languages in contact*. The Hague: Mouton.

Wesche, M., & Paribakht, T. S. (2010). *Lexical inferencing in a first and second language: Cross linguistic dimensions*. Bristol: Multilingual Matters.

White, J. L., & Horst, M. (2012). Cognate awareness-raising in late childhood: Teachable and useful. *Language Awareness*, *21*(1-2), 181–196.

Williams, J. (1995). Focus on form in communicative language teaching: Research findings and the classroom teacher. *TESOL Journal*, *4*(4), 12–16.

Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics*, *19*(3), 295–333.

Zietek, A. A., & Roehr, K. (2011). Metalinguistic knowledge and cognitive style in Polish classroom learners of English. *System*, *39*, 417–426.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 5–19). Cambridge: Cambridge University Press.

Profil Linguistique

Nom		Prénom		
Collège		Âge		F ☐ M ☐
Programme				

1. **Le <u>français</u> est ma**
    ☐ 1re langue   ☐ 2e langue   ☐ 3e langue   ☐ 4e langue   ☐ 5e langue

    OU

    ☐ J'ai appris le français et _____ (une autre langue) durant mon enfance.

2. **L'<u>anglais</u> est ma**
    ☐ 1re langue   ☐ 2e langue   ☐ 3^e langue   ☐ 4^e langue   ☐ 5^e langue

    OU

    ☐ J'ai appris l'anglais et _____ (une autre langue) durant mon enfance.

3. **L'<u>espagnol</u> est ma**
    ☐ 1re langue   ☐ 2e langue   ☐ 3e langue   ☐ 4e langue   ☐ 5e langue

4. **L'<u>allemand</u> est ma**
    ☐ 1re langue   ☐ 2e langue   ☐ 3e langue   ☐ 4e langue   ☐ 5e langue

5. **Langue(s) parlée(s) à la maison**
    Il est possible d'indiquer plus d'une langue pour chaque situation.

    a. Avec ma mère, je parle   _____

    b. Avec mon père, je parle   _____

    c. Avec mes grands-parents, je parle   _____

    d. Avec mes frère(s) et sœur(s), je parle   _____

6. **Veuillez indiquer l'âge auquel vous avez commencé à apprendre l'anglais, dans le type de cours d'anglais** *(au niveau primaire)* **qui s'applique à vous :**

   ☐ **Anglais régulier :** _____ ans   ☐ **Anglais intensif :** _____ ans

7. **Veuillez indiquer le nombre d'années scolaires complètes OU le nombre approximatif d'heures de cours (si moins d'une année scolaire) durant lesquelles vous avez reçu de l'enseignement de l'anglais.** *Additionnez le nombre total d'années et/ou d'heures pour chaque type de cours* *depuis votre entrée à l'école secondaire.*

   ☐ **Anglais régulier :** _____ années scolaires   **OU** _____ heures de cours

   ☐ **Anglais enrichi:** _____ années scolaires   **OU** _____ heures de cours

8. **Veuillez indiquer le nombre d'années, de mois et/ou de semaines (s'il y a lieu) pendant lesquels vous avez vécu dans les environnements linguistiques suivants.**

	Années	Mois	Semaines
Un pays ou une province où l'on parle l'anglais			
Une famille où l'on parle l'anglais			
Un lieu de travail où l'on parle l'anglais			
Une école anglophone			

9. **Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la langue anglaise:**

   ☐ très faible   ☐ faible   ☐ moyen   ☐ fort   ☐ très fort

10. **Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la/les culture(s) (coutumes, histoire, gastronomie, arts, architecture, littérature, etc.) anglophone(s) :**

   ☐ très faible   ☐ faible   ☐ moyen   ☐ fort   ☐ très fort

Concernant votre connaissance de l'<u>espagnol</u>

11. Veuillez indiquer l'âge auquel vous avez commencé à apprendre l'espagnol : _____ ans

12. Veuillez indiquer le nombre d'années scolaires complètes OU le nombre approximatif d'heures de cours (si moins d'une année scolaire) durant lesquelles vous avez reçu de l'enseignement de l'espagnol.

_____ années scolaires   **OU**   _____ heures de cours

13. Veuillez indiquer le nombre d'années, de mois et/ou de semaines (s'il y a lieu) pendant lesquels vous avez vécu dans les environnements linguistiques suivants.

	Années	Mois	Semaines
Un pays ou une province où l'on parle l'espagnol			
Une famille où l'on parle l'espagnol			
Un lieu de travail où l'on parle l'espagnol			
Une école hispanophone			

14. Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la langue espagnole:

☐ très faible ☐ faible ☐ moyen ☐ fort ☐ très fort

15. Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la/les culture(s) (coutumes, histoire, gastronomie, arts, architecture, littérature, etc.) <u>hispanophone</u>(s):

☐ très faible ☐ faible ☐ moyen ☐ fort ☐ très fort

16. Veuillez indiquer l'âge auquel vous avez commencé à apprendre l'allemand : _____ ans

17. Veuillez indiquer le nombre d'années scolaires complètes OU le nombre approximatif d'heures de cours (si moins d'une année scolaire) durant lesquelles vous avez reçu de l'enseignement de l'allemand.

_____ années scolaires **OU** _____ heures de cours

18. Veuillez indiquer le nombre d'années, de mois et/ou de semaines (s'il y a lieu) pendant lesquels vous avez vécu dans les environnements linguistiques suivants.

	Années	Mois	Semaines
Un pays où l'on parle l'allemand			
Une famille où l'on parle l'allemand			
Un lieu de travail où l'on parle l'allemand			
Une école germanophone			

19. Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la langue allemande:

☐ très faible ☐ faible ☐ moyen ☐ fort ☐ très fort

20. Sur l'échelle suivante, veuillez indiquer votre niveau d'intérêt à l'égard de la/les culture(s) (coutumes, histoire, gastronomie, arts, architecture, littérature, etc.) germanophone(s):

☐ très faible ☐ faible ☐ moyen ☐ fort ☐ très fort

21. Veuillez indiquer la fréquence avec laquelle vous vous servez du français, de l'anglais, de l'espagnol et/ou de l'allemand dans diverses activités. Entourez la lettre qui correspond à la fréquence de l'activité (voir en bas) pour chaque langue en question. Ajoutez d'autres activités s'il y a lieu.

Activités	En français	En anglais	En espagnol	En allemand
1. Interagir (oralement) avec des amis	Q S M A J	Q S M A J	Q S M A J	Q S M A J
2. Interagir (oralement) avec la famille	Q S M A J	Q S M A J	Q S M A J	Q S M A J
3. Ecouter la radio	Q S M A J	Q S M A J	Q S M A J	Q S M A J
4. Lire des journaux, revues, etc.	Q S M A J	Q S M A J	Q S M A J	Q S M A J
5. Lire des livres	Q S M A J	Q S M A J	Q S M A J	Q S M A J
6. Ecrire des lettres, cartes, courriels, etc. (personnels)	Q S M A J	Q S M A J	Q S M A J	Q S M A J
7. Ecrire des lettres, courriels (formels)	Q S M A J	Q S M A J	Q S M A J	Q S M A J
8. Ecrire des textos, clavarder sur internet, etc.	Q S M A J	Q S M A J	Q S M A J	Q S M A J
9. Regarder des émissions, films, séries (télé, cinéma, internet)	Q S M A J	Q S M A J	Q S M A J	Q S M A J
10. Jouer à des jeux vidéo	Q S M A J	Q S M A J	Q S M A J	Q S M A J
	Q S M A J	Q S M A J	Q S M A J	Q S M A J
	Q S M A J	Q S M A J	Q S M A J	Q S M A J
	Q S M A J	Q S M A J	Q S M A J	Q S M A J

Q=Quotidiennement; S=Quelques fois par semaine; M=Quelques fois par mois; A=Quelques fois par an; J=Jamais

Appendix B. *Metalinguistic ability test*

<div style="border:1px solid black; background:#bbbbbb; padding:10px; text-align:center;">

Test d'habiletés métalinguistiques
THAM – 3

</div>

Ce test comprend 3 parties : 1. Compréhension, 2. Langage figuré, 3. Acceptabilité.
Vous disposerez de 45 minutes pour compléter les parties 1 et 2 de ce test, dont la première partie comporte trois (3) questions tandis que la deuxième n'en comporte qu'une (1) seule. Après avoir lu les consignes et les exemples pour chacune des questions, prenez un maximum de 10 minutes pour répondre à chacune d'entre elles afin de permettre un bon déroulement du test. Cependant, le temps alloué à chacune des questions ne constitue pas un critère d'évaluation. La troisième partie du test prendra 45 minutes et aura lieu à la 2e période. Vous pourrez prendre votre pause habituelle avant d'entamer la dernière partie du test.

BONNE CHANCE !

<div style="border:1px solid black; background:#bbbbbb; padding:10px; text-align:center;">

1. Partie : Compréhension
(3 questions)

</div>

1.1 Relations qualitatives

Les phrases suivantes comportent des relations qualitatives. Indiquez s'il s'agit du même type de qualité dans les deux phrases et expliquez ce qui vous permet de répondre ainsi.

Exemple :

- Il a apporté une solution acceptable au problème.

- Il a apporté une bonne solution au problème.

Réponse :

Il ne s'agit pas du même type de qualité. Les deux phrases sont différentes.
Les adjectifs «acceptable» et «bonne» présentent des nuances. Dans la première phrase, «il a trouvé une solution acceptable» veut dire qu'il a trouvé une solution qu'on ne refuse pas. Elle est passable. Cela suggère qu'elle pourrait être mieux. Elle est acceptable pour les gens concernés, par exemple. Par contre, dans la deuxième phrase, «il a trouvé une bonne solution» veut dire que la solution est plus qu'acceptable, elle est «bonne». La solution est donc de qualité et présente des avantages.

1.1 a)

- **Il a une volonté de fer.**
- **Il est borné.**

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

1.2 Relations temporelles

Les phrases suivantes comportent des relations de temps entre deux événements. Indiquez si l'on trouve la même relation de temps dans les deux phrases et ce qui vous permet de répondre ainsi.

<u>Exemple</u> :

- Lisez d'abord, réfléchissez ensuite.
- Commencez à réfléchir après avoir lu.

Réponse :

Les relations temporelles de ces deux phrases sont différentes.
Dans la première phrase, il faut lire en premier et réfléchir après. Il y a une gradation temporelle, l'action de lire précède l'action de réfléchir. Dans la deuxième phrase, on comprend que la lecture vient aussi en premier lieu. Toutefois, la réflexion doit débuter après la lecture (*commencez à réfléchir*). L'accent est mis sur le moment du début de la réflexion. Dans la première phrase, rien n'indique que la réflexion n'ait pas commencé avant la lecture. De plus, alors que dans la première phrase, on a une suite réelle d'actions: 1. lire, 2. réfléchir; dans la deuxième, bien que l'ordre des actions soit différent au niveau graphique : 1. réfléchir, 2. lire, l'ordre des actions à entreprendre est généralement le même: la lecture précède la réflexion. La différence est au niveau de l'accent mis sur chaque action: lire dans la première phrase (*lisez d'abord*) et réfléchir dans la deuxième (*commencez à réfléchir*).

1.2 a)

- **Avant tout, soyez prudent.**
- **Avant d'écrire, pensez à ce que vous voulez dire.**

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

1.3 Test morphologique

Dans les deux phrases suivantes, le mot « un » est utilisé. À votre avis, les deux « un », jouent-ils le même rôle ? Expliquez en quoi ce rôle est identique ou différent.

<u>Exemple</u> :

- J'ai lu un poème de Beaudelaire que je ne connaissais pas.
- On m'a donné trois feuilles blanches et un crayon pour faire cet exercice.

Réponse :

Les deux « un » ne jouent pas le même rôle même si dans les deux phrases ils évoquent une idée d'unité. Dans la première phrase, le « un » est un article indéfini. Il désigne un des poèmes de Beaudelaire d'une manière indéterminée. Le « un » dans cette phrase n'est pas un nombre. Par contre, dans la deuxième phrase, le « un » est un nombre, un adjectif numéral. Je n'ai pas eu deux ou trois crayons, mais un seul. Le « un » indique la quantité.

1.3 a)

- **J'ai deux amis anglophones : l'un vient d'Angleterre et l'autre des Etats-Unis.**
- **Un café et deux thés pour la table 6 !**

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

2. Partie : Compréhension du langage figuré
(1 question)

2.1 Slogans publicitaires

Vous trouverez ci-dessous un slogan publicitaire tiré d'une annonce publicitaire. L'objet auquel il fait référence est indiqué à côté du slogan, entre parenthèses. Indiquez comment ce slogan devrait être compris et justifiez l'utilisation des mots qui le composent.

2.1 a)

« N'eau fatigue, n'eau stress» (Publicité pour une eau minérale)

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

3. Partie : Acceptabilité

Dans le texte suivant, vous trouverez un certain nombre de fautes. Indiquez-les, corrigez-les et justifiez chacune de vos corrections. Vous disposez de 45 minutes pour compléter cette partie.

« Colonel Fillmore est resté de son bureau et il a regardé par la fenêtre vers la direction nord le petit triangle de désert de lequel les falaises ne cachaient pas et vit une bande de points noirs qui se déplaceraient comme des fourmis tout droit en sa direction, vers la Forteresse, et elles commençaient à ressembler à des soldats immédiats.

De temps en temps, un officier était entré, quelques fois le lieutenant-colonel Dubuc ou le capitaine des inspections ou encore des officiers dans leur fonction.

Ils entraient sous différents prétextes, attendant impatiemment des ordres, le rapportant des nouvelles insignifiantes : qu'un nouvel arrivage de provisions alimentaires est livré dedans la ville; que les travaux de réparation des fours commençaient ce matin-ci; qu'environ dix soldats en congé devront déjà être de retour; que le télescope est installé sur le toit de forteresse centrale, si jamais le colonel choisit d'en tirer profit. »

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Test items in sentence context

1. Apfel

   Möchtest du einen *Apfel*?

2. Feld

   Die Kinder laufen über das *Feld*.

3. Fisch

   Ich mag keinen *Fisch*.

4. Freund

   Peter ist mein *Freund*.

5. Sohn

   Wir haben einen *Sohn*.

6. Schiff

   Ein *Schiff* fährt über den Fluss.

7. Blut

   Sie hat zu viel *Blut* verloren!

8. Berg

   Das ist ein großer *Berg*.

9. Herz

   Martina hat ein gutes *Herz*.

10. Meer

    Wir machen immer Urlaub am *Meer*.

11. Wasser

    Wasser ist lebenswichtig.

12. Ferien

    Dieses Jahr haben wir nur fünf
    Wochen *Ferien*.

13. Möbel

    Das sind alte *Möbel* aus Holz.

14. Kirche

    An Weihnachten gehen wir immer in
    die *Kirche*.

15. Messer

    Mit einem *Messer* kann man sich
    schneiden.

16. Haar

    Sabrina hat schwarzes lockiges *Haar*.

17. Fenster

    Von diesem *Fenster* kann man den
    Strand sehen.

18. Licht

    Bitte machen Sie das *Licht* aus!

19. Brust

    Mir tut die rechte *Brust* weh.

20. Wetter

    Seit einem Monat haben wir
    schlechtes *Wetter*.

21. Lied

    Wir hören zusammen ein *Lied*.

22. Zeitung

    Jeden Tag lesen wir die *Zeitung*.

23. Kleid

    Dei Frau trägt ein schönes *Kleid*.

24. Maus

    Die *Maus* ist ein kleines Nagetier.

25. Grund

Schwere Gegenstände sinken bis auf den *Grund*.

26. Milch

Viele Kinder lieben *Milch*?

27. Krawatte

Der Herr trägt eine bunte *Krawatte*.

28. Buch

Ich kaufe mir ein spannendes *Buch*.

29. Luft

Hier ist die *Luft* gut.

30. Sturm

Morgen gibt es einen *Sturm*.

31. Kuh

Die *Kuh* steht auf der Weide.

32. Socke

Ich finde die zweite *Socke* nicht.

33. Knie

Nach dem Laufen tut mein *Knie* weh.

34. Pferd

Das ist ein kräftiges *Pferd*.

35. Nachbar

Stefan ist mein netter *Nachbar*.

36. Gast

Herr Müller ist ein seltener *Gast*.

37. Sonne

Die *Sonne* geht auf.

38. Stein

Dieser Boden ist aus *Stein*.

39. Bäckerei

Das ist die beste *Bäckerei* der Stadt.

40. Spiegel

Ich schminke mich immer vor diesem *Spiegel*.

41. Tasse

Das ist eine schöne *Tasse*.

42. Regen

Eine Woche lang gab es nur *Regen*.

Bonjour et bienvenue!

Ceci est une tâche de traduction.

Vous allez traduire des mots de l'allemand vers le français.

Il y en a 42 au total et ce sont tous des noms.

C'est normal que la plupart des mots vous soient inconnus. Mais ne vous découragez pas! Essayez de faire des liens avec toutes vos autres connaissances pour en déduire le sens.

Chaque mot sera présenté sur l'écran, d'abord de façon isolée et ensuite dans le contexte d'une phrase. Vous entendrez la prononciation de chaque mot lorsque ce dernier apparaîtra sur l'écran.

Dès que le mot apparaît, vous avez 30 secondes pour en deviner la traduction et la dire à voix haute.

Pendant que vous êtes en train de chercher la bonne traduction, SVP, pensez tout haut, c'est-à-dire prononcez à voix haute toutes les choses qui vous viennent à l'esprit!

Même quand un mot ne vous inspire rien et qu'aucune traduction possible ne vous vient à l'esprit, dites-le également… Vos réponses seront enregistrées automatiquement.

Cliquez sur la touche ESPACE pour commenc

283

Appendix E. *Coding of the think-aloud protocols (TAPs)*

Code	Label	Description	Examples of correct translations	Examples of erroneous translations
PK	Prior knowledge	• of translation equivalent w/o statement of resemblance	• *I've already learned that word in class (or elsewhere)* • *Well, that's … I know that*	
A	Absence	• of translation • of reasoning	• *I don't know / have no idea…* → also: too many alternatives without any reasoning	
I	Intuition	a) of correct translation → based on English cognate, but no explicit reference b) random clang association	*I think it's […] / it could/might be […] / maybe […] or […] or […]* (spanning both columns) • *« Ça me fait penser à souris, Maus… je sais pas »* • *« J'sais pas c'est quoi, ça me dit rien… Wetter… euh... peut-être température, j'sais pas… »*	• clang association with a word from a different word class (e.g., Blut: *blue*, Feld: *felt*, etc.) • reference to a noun that <u>could not possibly fit the context</u> (e.g., *gehen* + "quiche", "chiffre"/"shift" + *fahren*, etc.) • enumeration of <u>several options</u> based on clang association
N	Noticing	a) of formal resemblance to English cognate b) of formal resemblance to a different noun (English or other)	*It looks / sounds like / makes me think of…* (spanning both columns) • *« Ça me fait penser à cœur, genre heart… Herz […] »* • *« J'ai, eum… church, l'église… Kirche […] ça me fait penser à ça, église »*	• reference to a noun that could actually fit the context, i.e., no counter-indications from sentence or internal logic (e.g., *brosse, crèche, knife, blouse, water* for *"Wetter"*) • random clang association + noticing of inappropriacy
MA	Meta-awareness	<u>**Report of line of thought**</u> • orthographic/phonetic resemblance explicitly stated (relation between levels or specific letters/sounds) • reasoning based on internal logic • cross-check with sentence (or other learning events) • reference to typological proximity/distance	*Justifications of the following types* (spanning both columns) • *it's sounds the same, even though it's not spelled the same…* • *it's just the /i/ that is missing, otherwise it's the same…* • *it makes sense when I look at the sentence…* → also: pronunciation/translation of sentence = confirmation • *English and German are often very close*	• *"nach" is like night and "bar" must be a bar, so it's probably a nightclub* • *"stein" must be a place, many cities end in "stein"* • *but it must be a noun, it's spelled with a capital letter* • *but it doesn't make sense with the rest of the sentence* → also when sentence is analyzed yet leading to an incorrect translation • *maybe it's derived from English or the other way around*
UR	Underlying rules	**Report of underlying rules or structures and/or related meanings** • phonetic/orthographic • syntactic/semantic	• *the /p/ sound often becomes /f/ in German* • *"Sohn" and "Sonne", must not be confused, they look very similar, but they don't sound the same, I have to remember that…* → here: explicit reference to another word that they could have confused the target item with… but explicit statement of how they arrive at keeping them distinct • *"finde" looks like "find", so by deduction it must be a sock "I cannot find the sock"* → here, analysis co-occuring word for meaning & function = confirmation • *"Gast" must be "guest" and "seltener" that must be the adjective that goes with it* → here: thorough analysis of the whole sentence and confirmation based on meaning & function of co-occurring words	

Appendix F. *Michigan Test of English Language Proficiency (MTELP)*

FORM **R**

MICHIGAN TEST OF ENGLISH LANGUAGE PROFICIENCY

PREPARED BY A. CORRIGAN, B. DOBSON, E. KELLMAN, M. SPAAN, L. STROWE, AND S. TYMA.

The Michigan Test of English Language Proficiency (MTELP), a test of grammar, vocabulary, and reading, was used as one part of the official Michigan Test Battery by the English Language Institute Testing Service of The University of Michigan. At the time this form of the MTELP was used by the ELI Testing Service, a Michigan Test Battery consisted of a written composition, a listening test, and an MTELP.

INSTRUCTIONS: Do not begin the test until the examiner has read these instructions with you.

1. This test contains 100 items. These are divided into three parts: Part I is a grammar test;
Part II is a vocabulary test; and Part III is a test of reading comprehension. There are ex-
amples at the beginning of each part to show you how to answer the problems.

2. Each problem in the test has only one best answer. Make only one mark on your answer sheet
for each problem. If you want to change an answer, erase the old mark or cross it out several
times.

3. Do not make any marks on this test booklet. Write only on your answer sheet.

4. The examiner will not explain any test problem; only the examples in each part may be explained
if you do not understand how to do the problems.

5. Try to answer all the problems. Each problem counts the same.

6. You will have 75 minutes to finish the entire test. Do not spend too much time on any one
problem or you will not have time to finish the test. If you do not know the answer to a prob-
lem, you should make a reasonable guess. Unanswered problems will be counted wrong.

7. Continue working until you have finished the test. Do not stop after each part.

8. Begin the test now.

1

286

PART I

GRAMMAR

This is a test to show how well you can recognize and use English grammatical structures. Each question in this test is part of a conversation. In each conversation a word (or group of words) is left out. Following the conversation are four choices of words which might be used in the incomplete conversation. You are to select the word (or group of words) which would be used by a speaker of English, and which will best fit into the conversation.

EXAMPLE A:

"What is that thing?"
"That _____ a spider."

    a) to call
    b) for calling
    c) be called
    d) is called

The correct English sentence is: "That is called a spider."

To show that d, is called, is the best answer to this example a cross has been made next to d for Example A on the answer sheet.

Answer all the questions of Part I in this manner.
Mark only one answer for each problem.

CONTINUE

1. "Would you like more rice?"
   "No, I have already had _____."

    a) to eat enough
    b) enough to eat
    c) enough eat
    d) eat enough

2. "How old is Van?"
   "I think he's _____ I am."

    a) the same old as
    b) as age as
    c) as same age as
    d) the same age as

3. "Why couldn't you play football last Saturday?"
   "Because my wife made me _____ all the floors in the house."

    a) washed
    b) wash
    c) washing
    d) to wash

4. "Hurry up with those suitcases. We're late for the train!"
   "But all these suitcases are _____ for me to carry."

    a) heavy too much
    b) much too heavy
    c) much heavy
    d) too much heavy

5. "Is everyone here?"
   "Yes, everyone _____ Paul."

    a) but
    b) not
    c) than
    d) without

6. "Is she sorry she left home?"
   "She doesn't seem to regret _____ at all."

    a) had left
    b) to leave
    c) left
    d) having left

2

287

7. "Does Eleanor grow roses?"
"Yes, but she grows many other _____, too."

   a) kinds of flower
   b) kinds of flowers
   c) kind flowers
   d) kind of flowers

8. "Couldn't you find a place to stay in downtown Toronto?"
"No, we _____ reservations before we went."

   a) have made
   b) should had made
   c) were made
   d) should have made

9. "Bad weather never bothers Sheila."
"Yes. She is always happy, _____ what the weather is like."

   a) which doesn't matter
   b) doesn't matter
   c) however
   d) no matter

10. "That's a beautiful ring."
"I'm only borrowing it. It is _____."

    a) of my mother
    b) my mother's
    c) belong to my mother
    d) belonging to my mother

11. "Do you like apple pie?"
"Yes, almost _____ I like peach pie."

    a) as much
    b) as much as
    c) so much as
    d) as

12. "You look tired today."
"I feel tired too. I wish I _____ to bed earlier last night."

    a) have gone
    b) were going
    c) had gone
    d) did go

13. "This car has very low mileage."
"It was owned by a little old lady who hardly _____ drove it."

    a) ever
    b) often
    c) sometimes
    d) never

14. "Do the police know where the bank robbers are?"
"No, but they are believed _____ somewhere in the city."

    a) to be hide
    b) be hiding
    c) to hide
    d) to be hiding

15. "Should we hire Jones or Taylor?"
"I don't care, but _____ way you decide, let me know."

    a) whichever
    b) some
    c) what
    d) however

16. "Do you still need help on that math problem?"
"No, Anne explained _____."

    a) me to many things
    b) many things me
    c) me many things
    d) many things to me

17. "The flowers you planted last month are growing fast."
"Yes, and they _____ leaves already."

    a) are starting get
    b) start getting
    c) are start to get
    d) are starting to get

18. "Why don't you like working for Vance?"
"He's a hard man _____."

    a) that I please
    b) for pleasing
    c) to please
    d) to be pleased

288

19. "What did you study in college?"
"I majored in English _____ under-graduate."

a) as
b) as an
c) in
d) in the

20. "Did you enjoy the lecture about the Amazon River basin?"
"Yes, we were told all about _____."

a) there how is life
b) how life is there
c) how is life there
d) life is how there

21. "Why didn't Emily come?"
"I don't know. She gave me _____ of excuses."

a) all sort
b) every sorts
c) all sorts
d) every sort

22. "No, I'm not Greta Garbo."
"You look _____ her that I thought you were."

a) as much like
b) so much like
c) as much as
d) like as much

23. "Joan couldn't go to the London Philharmonic concert."
"She could _____ it last night on television."

a) has seen
b) have seen
c) see
d) have been seen

24. "Can that be a flying saucer?"
"I don't know _____ ____."

a) what can
b) what it can
c) what is it
d) what it is

25. "I can't find my keys."
"Do you remember _____ them down anywhere?"

a) to have put
b) you put
c) to put
d) putting

26. "Does Debbie have to type much at her new job?"
"No, but _____ little typing she does is well done."

a) when
b) how
c) what
d) such

27. "You'll discover that my sister is always willing to help people."
"That _____ is enough to make me like her."

a) alone
b) only
c) single
d) lonely

28. "When do you expect Charles?"
"He _____ here any minute."

a) ought be
b) would be
c) should be
d) must be

29. "Was anyone hurt as a result of the hurricane?"
"No, but the high winds caused _____ to buildings."

a) many damages
b) many damage
c) a lot of damage
d) damage a lot

30. "What happened to Elizabeth?"
"She broke her leg _____ on the river last winter."

a) on skating
b) to skate
c) for skating
d) skating

289

31. "What is that?"
    "This machine _____ removing pits
    from cherries."

    a) for
    b) is
    c) is for
    d) is to

32. "He'll never succeed."
    "But he's determined _____."

    a) try
    b) trying
    c) to try
    d) on trying

33. "Is Heinrich a good worker?"
    "Yes, he _____ at the store for only
    two weeks when he was promoted."

    a) had been working
    b) has worked
    c) is working
    d) has been working

34. "You seem angry at Martha."
    "I am. _____ I'm concerned, she
    can go away forever."

    a) As like as
    b) As such as
    c) As much as
    d) As far as

35. "Have you and the children always gone
    for rides on Sunday afternoon?"
    "No, when they were younger, I _____
    take them for walks in the park."

    a) may
    b) had
    c) should
    d) would

36. "You don't look very cheerful."
    "I have just had a long, _____ day."

    a) exhausting
    b) exhaust
    c) exhausted
    d) exhaustible

37. "I smell something burning in the kitchen."
    "George must _____ to cook something."

    a) have been tried
    b) have been trying
    c) had tried
    d) be try

38. "Didn't you buy gifts while you were here?"
    "No, _____ was the scenery, not the
    stores."

    a) for what I went
    b) what I went for
    c) I went for
    d) what I went

39. "I have to stop at the store to get Ted a
    pack of cigarettes."
    "You shouldn't buy him cigarettes if you
    want _____ smoking."

    a) stop him
    b) that he stops
    c) him stop
    d) him to stop

40. "Does Susan resemble Sally?"
    "Yes, they look _____."

    a) like
    b) alike
    c) likely
    d) like the same

CONTINUE TO PART II

290

VOCABULARY

There are two types of vocabulary items in this test. In the first type you are given a sentence followed by four words or phrases. You are to find the word or phrase that is closest in meaning to the underlined word (or words) in the sentence and that could be used in the sentence without changing its meaning greatly.

EXAMPLE B: It's too windy to go for a stroll.

      a) swim
      b) sail
      c) drive
      d) walk

The word 'walk' means about the same thing as 'stroll' in this sentence. The sentence "It's too windy to go for a walk," means the same thing as "It's too windy to go for a stroll." To show that d, walk is the correct answer, a cross has been made in the space next to d for Example B on the answer sheet.

In the other type of item you are given a sentence with one word omitted and a list of four words. You are to find the word that would best complete the sentence.

EXAMPLE C: Because of the storm and rough waves, it would be foolish to go out sailing today in a small _____.

      a) automobile
      b) house
      c) boat
      d) beast

The word 'boat' fits best in the sentence so that it reads, "Because of the storm and rough waves, it would be foolish to go out sailing today in a small boat." To show that c, boat is the correct answer, a cross has been made in the space next to c for Example C on the answer sheet.

Answer all the questions of Part II in this manner.
Mark only one answer for each problem.

CONTINUE

41. He adhered to the rules.

    a) followed
    b) didn't obey
    c) justified
    d) was angered by

42. Since his car accident, he walks with a slight _____.

    a) clip
    b) bump
    c) whimper
    d) limp

43. Jeff tried to conceal the truth.

    a) find
    b) accept
    c) hide
    d) learn

44. The boys wanted to sail down the river, so they tied logs together to make a _____.

    a) shoal
    b) prop
    c) raft
    d) moat

6

45. He works under <u>adverse</u> conditions.

   a) unusual
   b) normal
   c) changing
   d) difficult

46. Fred always goes to the doctor to complain about some imaginary _____.

   a) ailment
   b) sanitation
   c) irrigation
   d) filament

47. Charles walked along the <u>crest</u>.

   a) bottom
   b) top
   c) beach
   d) path

48. 4000 people signed a _____ demanding that the mayor resign from office.

   a) petition
   b) lease
   c) profile
   d) repose

49. The cake <u>crumbled</u>.

   a) fell apart
   b) was eaten
   c) was burnt
   d) spoiled

50. He tried to _____ his tears when he heard his brother had been killed.

   a) check on
   b) clear out
   c) hold back
   d) force down

51. That custom is <u>prevalent</u> here.

   a) uncommon
   b) wide-spread
   c) old-fashioned
   d) not legal

52. He <u>hoarded</u> his money.

   a) lost
   b) kept
   c) spent
   d) gave away

53. Stan became _____ to an experienced electrician so he could learn the trade.

   a) subscribed
   b) enrolled
   c) installed
   d) apprenticed

54. He did not expect any <u>mishaps</u>.

   a) protests
   b) visitors
   c) accidents
   d) adventures

55. I'm afraid the car may break down; it's been making _____ noises.

   a) durable
   b) sordid
   c) ominous
   d) wanton

56. No one was able to <u>console</u> Susan.

   a) accompany
   b) comfort
   c) control
   d) understand

57. Nowadays there is _____ shortage of food in many parts of the world.

   a) an acute
   b) a barren
   c) a fertile
   d) an elaborate

58. He was unhappy <u>prior to</u> reading the letter.

   a) before
   b) after
   c) in spite of
   d) because of

292

59. We were perplexed by his remarks.

    a) amused
    b) confused
    c) fooled
    d) upset

60. He doesn't look like his mother, but he _____ his father.

    a) takes over
    b) takes on
    c) takes after
    d) takes up

61. The director told us to stay out of the controversy.

    a) argument
    b) program
    c) garden
    d) meeting

62. Whenever Steve wears a new tie, he _____ something on it.

    a) winks
    b) bruises
    c) spills
    d) scowls

63. They slaughtered the pig.

    a) fattened
    b) cooked
    c) ate
    d) killed

64. This information should not be given to anyone; it is to be kept _____.

    a) venerable
    b) confidential
    c) latent
    d) scant

65. Their supplies dwindled.

    a) were needed
    b) were stolen
    c) decreased
    d) increased

66. His note was dispatched.

    a) refused
    b) lost
    c) sent
    d) destroyed

67. Andy's sudden decision to quit school made a _____ change in his life.

    a) drastic
    b) blunt
    c) ponderous
    d) strenuous

68. He did not use the right utensil.

    a) information
    b) tool
    c) method
    d) amount

69. When Harry and his wife began to argue, her mother promised not to _____.

    a) meddle
    b) insert
    c) impart
    d) divert

70. He reveres his teacher.

    a) avoids
    b) dislikes
    c) respects
    d) trusts

71. They didn't want to pay tax on the goods, so they _____ them over the border.

    a) contrived
    b) migrated
    c) alluded
    d) smuggled

72. I could tell by looking at her that she was taken aback.

    a) worried
    b) tired
    c) shocked
    d) disappointed

73. The courses listed here are <u>compulsory</u>.

    a) practical
    b) required
    c) hard
    d) interesting

74. Last year I spent too much money so I decided to _____ my income carefully this year.

    a) budget
    b) enlist
    c) schedule
    d) ransom

75. He did it <u>deliberately</u>.

    a) unwillingly
    b) unexpectedly
    c) on purpose
    d) democratically

76. Many of these laws are very old-fashioned; they should be _____ and replaced by new ones.

    a) decreed
    b) recruited
    c) uplifted
    d) abolished

77. I walk so slowly that even though I had a _____, they got there before me.

    a) head start
    b) mad dash
    c) long shot
    d) running foot

78. Their work was <u>arduous</u>.

    a) profitable
    b) exciting
    c) difficult
    d) beneficial

79. His good qualities were _____ by his bad ones.

    a) detached
    b) interposed
    c) offset
    d) reproved

80. When we have the <u>inclination</u>, we'll visit Ellen.

    a) invitation
    b) address
    c) time
    d) desire

CONTINUE TO PART III

READING COMPREHENSION

This is a test to show how well you read English. There are several reading passages each followed by some questions about the passage. You should read each passage carefully and then try to answer the questions following that passage. If you do not know the answer at first, you may read the passage again, but <u>do not spend too much time on the passage</u> or you will not have enough time to finish.

EXAMPLE:

While I was getting ready to go to town one morning last week, my wife handed me a little piece of red cloth and asked if I would have time during the day to buy her two yards of cloth like that. I told her I would be glad to do it. And putting the piece of cloth into my pocket, I took the train for town.

D. The person telling the story is . . .

      a) a married lady.
      b) an unmarried lady.
      c) a married man.
      d) an unmarried man.

You know that the person telling this story is <u>a married man</u> because he says, ". . . my wife handed me . . ." Because <u>c</u>, <u>a married man</u> is the correct answer, a cross has been made in the space next to <u>c</u> for Example D on the answer sheet.

E. When was the author given a piece of red cloth?

      a) In the morning.
      b) At noon.
      c) In the afternoon.
      d) In the evening.

The passage says, ". . . one morning last week, my wife handed me a little piece of red cloth . . ." To show that <u>a</u>, <u>In the morning</u> is the correct answer, a cross has been made in the space next to <u>a</u> for Example E on the answer sheet.

Answer all questions of Part III in this manner.
Mark only one answer for each problem.

CONTINUE

Persecuted in Russia for their pacifism and their religious beliefs, the Doukhobors emigrated to Canada in search of freedom from oppression. When they arrived, in 1898, most of them were extremely poor. Their able-bodied men worked outside the farming community for wages while the women cleared the land, planted crops and gardens, built homes, took care of the livestock and attended to the regular household chores. In order to break the land for crops, the women hitched themselves to ploughs. Many people were horrified by accounts of women ploughing, but to the Doukhobors the courage of the women was an important symbolic event in their history, woven into their mystique of peace and useful toil.

Doukhobors viewed education with suspicion, fearing it would include patriotic and militaristic teachings. In British Columbia they refused to send their children to school, resulting in the removal of several hundred children from their homes without parental consent or court order. The Sons of Freedom sect protested by burning their homes and holding marches and nude demonstrations. When a hundred Doukhobors were jailed, Florence Storgoff led a Great March from Krestova to the Montail Prison in protest.

81. What did the Doukhobors hope to find in Canada?

    a) Freedom of religious belief.
    b) Better education for their children.
    c) Better farm land.
    d) Better jobs.

82. One prominent belief of the Doukhobors was that . . .

    a) useful work was important.
    b) children should always obey their parents.
    c) hard work would make them rich.
    d) women were inferior to men.

83. How did the Doukhobor men feel about their women ploughing?

    a) They felt it was a necessary evil.
    b) They were mystified by it.
    c) They had no strong feelings about it.
    d) They felt it was symbolically important.

84. Why were the children taken from their homes?

    a) Their homes had been burnt.
    b) They were members of the Sons of Freedom.
    c) They had demonstrated against the government.
    d) Their parents hadn't sent them to School.

85. What was the Great March?

    a) The emigration of Doukhobors from Russia to Canada.
    b) A protest against the jailing of the Doukhobors.
    c) A protest against the burning of Doukhobor homes.
    d) A protest against the Doukhobors.

296

Artemisia Gentileschi was a painter active in the first half of the seventeenth century, contemporary with Rubens and Bernini. She adopted the style of the influential Caravaggio, producing paintings of a forceful realism, which combined a penetrating observation of life and a dramatic lighting that accentuated volumes and textures. Most art historians believe that there is seldom if ever, any demonstrable quality of style or expression in painting by which the sex of the painter can be deduced, although the subject matter might be revealing. Artemisia seldom painted scenes from the New Testament which was unusual for a seventeenth century painter. A woman of unimpeachable virtue might have received commissions to paint religious subjects for the Church. However, she was not believed to be such a woman. Stories were circulated about her, though at least some of them are known to be false. She was an independent career woman with responsibilities not only to herself but also to a child whom she had borne during an early, unsuccessful marriage to a man named Stiattesi. There is no documentary evidence that patrons believed there ought to be restrictions on the subjects Artemisia might paint. But in nearly every one of her extant works a female is the principal subject, and usually a female under duress. Artemisia probably devoted most of her time to certain themes not for emotional reasons connected with her sex or her past but because they are what her male patrons ordered. In this respect a double standard was applied against her. The private lives of male painters, however dissolute—Caravaggio, for example—seldom seem to have affected the nature of the commissions they received.

86. Artemisia's style can be described as . . .

    a) light and textural.
    b) highly personal.
    c) realistic and dramatic.
    d) almost masculine.

87. Most of Artemisia's works involved . . .

    a) women with independent personalities.
    b) scenes from nature.
    c) scenes from the New Testament.
    d) women under duress.

88. The subjects Artemisia painted were influenced mostly by . . .

    a) the styles of the seventeenth century.
    b) what her patrons wanted from her.
    c) what the Church allowed.
    d) her femininity.

89. Artemisia's contemporaries apparently thought of her as . . .

    a) a woman of little virtue.
    b) an influential painter.
    c) an over-emotional painter.
    d) a woman of great virtue.

90. Commissions for church art work were not given to . . .

    a) any persons of doubtful virtue.
    b) any women.
    c) men of doubtful virtue.
    d) women of doubtful virtue.

In the last half century concepts of conservation have grown more sophisticated, less parsimonious. Natural conditions are no longer automatically presumed to be best. It has been demonstrated, for example, that controlled burning protects Southern pine woods better than absolute prohibition of fire, and that the shooting of doe may be necessary to prevent the starvation and extinction of deer. Resource managers have abandoned the goal of preservation, with its romantic concomitants, in favor of more flexible <u>present</u> use and development; and they distinguish among resources that are inexhaustible, those that are renewable, and those that are neither. But the public at large continues to hold to the older conservation mystique, according to which nature is always right and man is usually wrong.

91.   The author says that modern conservationists are mainly concerned with . . .

    a) prohibition.
    b) use and development.
    c) demonstration.
    d) romantic ideals.

92.   The author states that controlled burning . . .

    a) is a protective measure.
    b) is primarily a demonstration device.
    c) is destructive.
    d) prevents deer from starving.

93.   The goal of preservation was abandoned after changes were dictated by . . .

    a) powerful private groups.
    b) a less simple minded view of conservation.
    c) a reaction against romanticism.
    d) growing public concern about conservation.

94.   The author cites the case of deer to show that . . .

    a) deer is a renewable resource.
    b) there are too many deer now.
    c) deer are becoming extinct.
    d) the use of resources can actually preserve them.

95.   The author says that in matters of conservation the public is . . .

    a) more concerned than ever.
    b) demanding rapid changes.
    c) fifty years behind the times.
    d) not interested.

The American Indian voices of the past have not been lost. A few authentic accounts of American western history were recorded by Indians either in pictographs or in broken English, and some managed to get published in obscure journals, pamphlets, or books of small circulation. In the late nineteenth century, when the white man's curiosity about Indian survivors of the wars reached a high point, enterprising newspaper reporters frequently interviewed warriors and chiefs and gave them an opportunity to express their opinions on what was happening in the West. The quality of these interviews varied greatly, depending upon the abilities of the interpreters, and the inclination of the Indians to speak freely. Some feared reprisals for telling the truth, while others delighted in hoaxing reporters with tall tales and shaggy-dog stories. Contemporary newspaper statements by Indians must therefore be viewed with skepticism, although some of them are masterpieces of irony and others burn with outbursts of poetic fury. Among the richest sources of first person statements by Indians are the records of treaty councils and other formal meetings with civilian and military representatives of the United States government; recording clerks, using a new stenographic system, sat beside the official interpreter. Even when the meetings were in remote parts of the West, someone usually was available to write down the speeches, and because of the slowness of the translation process, most of what was said could be recorded in long-hand. Interpreters quite often were half-bloods who knew spoken languages but seldom could read or write. Like most oral peoples, they and the Indians depended upon imagery to express their thoughts, so that the English translations were filled with graphic similes and metaphors of the natural world. If an eloquent Indian had a poor interpreter, his words might have been made to seem flat or uninteresting, but a good interpreter could make a poor speaker sound poetic.

96. According to the passage, newspaper reporters interviewed Indians in order to . . .

    a) show how much imagery the Indians used.
    b) satisfy the white people's curiosity about Indians.
    c) provide an accurate historical record.
    d) create interest in Indian problems.

97. The white man's curiosity about the Indian was greatest . . .

    a) after 1950.
    b) just before 1800.
    c) just before 1900.
    d) between 1900 and 1950.

98. The author is skeptical of Indian statements in newspapers because . . .

    a) some Indians were afraid they would be punished if they told the truth.
    b) reporters had very little contact with Indians.
    c) the government usually regulated what the reporters could write about.
    d) the Indians used many metaphors.

99. The records of meetings between Indians and government officials are accurate records of what Indians said because the . . .

    a) interpreters were required to be half-bloods.
    b) slow translation process allowed accurate recording of what was said.
    c) Indians were very graphic in their expressions.
    d) Indians were always encouraged to speak freely at the meetings.

100. It is not always easy to tell if an Indian was eloquent because . . .

    a) poor interpreters could make eloquent Indians sound uninteresting.
    b) Indians were usually too afraid to speak.
    c) Indians usually spoke broken English.
    d) the English translations were usually poor.

END OF TEST

Check your answers. Be sure that your name and Form R have been written on your answer sheet.