In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
SERIA FILOLOGIA ANGIELSKA NR 48

Magdalena Wrembel

In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology

WYDAWNICTWO NAUKOWE

POZNAŃ 2015

ABSTRACT. Wrembel Magdalena. *In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology.* Adam Mickiewicz University Press. Poznań 2015. Seria Filologia Angielska nr 48. Pp. 430, 94 Figs., 190 Tabs. ISBN 978-83-232-2977-3. ISSN 0554-8144. Text in English with a summary in Polish.

The book explores the phenomenon of cross-linguistic influence (CLI) in phonological acquisition from a multilingual perspective. A distinction is made between the second vs. third language acquisition and the complexity of the latter process as well as its conditioning factors are elaborated on. The contribution aims to test the tenets of the current theoretical models of multilingual acquisition including the L2 Status Model, the Cumulative Enhancement Model and the Typological Primacy Model. To this end, three studies were conducted in parallel on four groups of participants with complementary language combinations (i.e. mirrored L2 and L3 sets). The studies involved (1) accentedness, comprehensibility and accuracy ratings assessing the perceived phonetic performance in the L3, (2) acoustic measurements of voice onset time (VOT) in the L1, L2 and L3 as a correlate of foreign accentedness, and (3) the degree of metaphonological awareness generated from oral protocols. The results were analysed separately for each study as well as globally by means of across groups and across studies comparisons. The findings indicate that CLI in the L3 phonological acquisition may have multiple sources including both the native and non-native languages, that it is gradual and structure dependent and that the proposed models can account only partially for its specificity.

KEY WORDS: third language (L3), phonological acquisition, cross-linguistic influence

Magdalena Wrembel, Faculty of English, Adam Mickiewicz University. al. Niepodległości 4, 61-678 Poznań, Poland; email: magdala@wa.amu.edu.pl

Reviewer/Recenzent: prof. dr Ulrike Gut

Cover design/Projekt okładki: Agnieszka Frydrychewicz
Cover design features a motif of “Abstrakcyjne dmuchawce” www.bazarek.pl/obrazy nowoczesne
Typsetting and formatting/Skład i formatowanie: WA UAM

ISBN 978-83-232-2977-3
ISSN 0554-8144

To my Mentor, Professor Katarzyna Dziubalska-Kołaczyk,
for her invaluable guidance and inspiration.

To my Family,
for their unfailing love and support.

Contents

List of Figures

List of Tables

Chapter 1

A new perspective: SLA vs. TLA

1.1. Introduction

Multilingualism has become a norm rather than exception in the contemporary world and a large part of the population speaks several languages on a daily basis. There is a growing recognition that it is a default state of human linguistic competence and that people possess a capacity to learn several languages. Some scholars even argue that in the present era of globalization, multilingualism has acquired a special significance at an unprecedented scale; thus they herald a new world order and a new linguistic disposition (Aronin and Singleton 2008). The phenomenon of multilingualism has become particularly relevant from the European perspective due to an increasing job-related mobility, a growing popularity of foreign exchange programmes, particularly in secondary and tertiary education, as well as the introduction of one or more foreign languages at the early stages of primary education, in accordance with the recommendations of the European Union (e.g. White Paper on Education and Training 1995).

The emergence of the complex linguistic landscape has raised many questions concerning language competence, its development, use, interaction or attrition. Consequently, investigations into the process of multiple language acquisition as well as the mutual influence of various language systems within a multilingual person's repertoire appear to be particularly valid and called for. A view that has recently been endorsed is that "research on multilingual behaviour can offer some valuable insights about the process of non-native language acquisition and speech production as a whole" (De Angelis 2007: 2). Limiting one's scope of inquiries to the second language cannot provide adequate information about language processing and use from a multilingual speakers' perspective. Prior linguistic knowledge and previous language learning experience have gained recognition as powerful factors in human cognition (e.g. Pennington 1999) leading to a growing understanding of the necessity to investigate the uniqueness and complexity of language acquisition beyond the first foreign language.

Another potential explanation of the current spread of multilingualism can be derived from a wealth of sociolinguistic factors, including sociolinguistic typology. In his recent book as well as earlier publications, Trudgill (2011) argues against common misunderstanding of complexity theories in linguistics claiming that the equicomplexity hypothesis has no validity. Although the human language learning faculty remains the same all over the world, changing sociolinguistic conditions lead to the loss of linguistic complexity under certain conditions such as the modern phenomenon of widespread adult acquisition. Through dealing with questions related to varying complexity of linguistic structures, Sampson et al. (2009) and Trudgill (2011) point to the interaction between acquisitional factors and social history. In an attempt to account for the processes of complexification and simplification of linguistic structures, Trudgill discusses social determinants such as the type of language contact. In the usual type of contact, the so called 'linguistic equilibrium', languages tend to maintain a normal level of complexity and this leads to stable bilingualism. Such a scenario applies mostly to smaller and more isolated languages or dialects (e.g. Islandic or Flemish), which are mostly acquired natively. The inherent characteristics of the so called 'societies of intimates' involving small size, dense social network and low contact, trigger further complexification (Trudgill 2011). For example, cross-linguistically marked phonological developments occur more frequently in isolated varieties rather than in widespread ones.

On the other hand, another type of language contact entails a rapid acquisition by a considerable number of adult learners. In this particular situation, due to somewhat diminished language learning abilities on the part of the late acquirers, grammars tend to become less elaborate, thus leading to simplification, as has been the case with English, for example. This is a typical scenario for larger and more complex human groups, the so called 'societies of strangers' with loose social networks but high adult contact situations. The process of linguistic simplification is reflected, among others, in the regularization of regularities, greater lexical and morphological transparency or loss of redundancy (cf. Trudgill 2011). When faced with a massive linguistic complexity typically associated with 'exotic' languages, the current moderate complexity of such widespread languages as English or French is considered to be a socio-historical anomaly.

> If widespread adult-only language contact is a mainly post-neolithic and indeed a mainly modern phenomenon associated with the last two thousand years, and if the development of large, fluid communities is also a post-neolithic and indeed mainly modern phenomenon, then according to this thesis the dominant standard modern languages in the world today are likely to be seriously atypical of how languages have been for nearly all of human history (Trudgill 2009: 109).

Therefore, the recent shift away from the equal complexity principle, both from the diachronic and synchronic perspectives, following from the aforementioned social determinants of linguistic simplicity and complexity, may be considered as a valid account of the current linguistic situation. The ensuing multilingual landscape is characterised by the predominance of some pluricentric languages, a drastic increase in high language contact situations and adult non-native acquisition.

Research into the acquisition of second language speech has enjoyed a well documented tradition; however, recently a new tendency has emerged in accordance with which several scholars have started to differentiate between the acquisition of the first foreign language (L2) as opposed to other subsequent languages (L3, L4, Ln). Consequently, Third Language Acquisition (TLA) has stated to be recognised as an independent field of inquiry and we can witness a dynamically growing body of related literature (e.g. Cenoz 2001, Cenoz et al. 2001, De Angelis 2007, Hufeisen 1994, Rothman et al. 2013, Sopata 2013). The increasing recognition of the spread of multilingualism, as well as the significance of language for the society, has resulted in a growing number of investigations into multilingual behaviour, primarily from the educational and sociolinguistic perspectives (e.g. Cenoz et al. 2001, Cummins 2001). In turn, the psycholinguistic and cognitive aspects of multilingualism have started to be the focus of research relatively late (see Rothman et al. 2013 for an overview).

Speech is a particularly relevant aspect of foreign language communication and it constitutes a multifaceted phenomenon with diverse sociolinguistic implications ranging from prestige to stigmatization, with pronunciation being a strong identity marker. Issues related to the mutual impact of several phonological systems and the intricacies of multilingual language acquisition appear to be particularly interesting and worth investigating. Therefore, the present contribution is intended to address the highly relevant question of the processes involved in multilingual speech production from a novel acquisition perspective.

1.2. From second to third language acquisition

At the initial stages of its development, research into multilingualism was regarded as the extension of well-established related research areas such as bilingualism and Second Language Acquisition. However, over time, the acquisition of a third language has gained recognition as being qualitatively different and the present section aims to discuss basic differences between Second and Third (or Additional) Language Acquisition, pointing to a monolingual bias in former studies on foreign language processing and production. Hammarberg (2009) explains this shift in focus in the research perspective by pointing out that scholars have started to differentiate between language learners on the basis of the complexity of their linguistic background (i.e. monolingual vs. bilingual). The main difference between the second and third language acquisition is that L3 learners have already acquired their first foreign language (i.e., L2), and thus they can rely on some conscious linguistic knowledge as well as language-learning experience and strategies (cf. Cenoz and Jessner 2000, De Angelis 2007).

Moreover, scholars working on this new perspective maintain that TLA is inherently more complex than SLA and it involves a quality change in the language learning and processing (Cenoz et al. 2001, Hufeisen and Lindemann 1997). One of the pioneers and ardent proponents of researching multilingualism and third language acquisition as a subfield of its own, pointed to differences inherent in psycholinguistic processing: "A comparison of bilingual and trilingual processing suggests that these similarities and differences are both of a quantitative and qualitative kind, and therefore trilingual competence is distinct from bilingual competence" (Hoffmann 2001: 1). This stance was seconded by Jessner (2006) who put forward a related argumentation: "[…] the process and the product of having learnt a second language can potentially exert influence on the acquisition of an L3 and this involves a quality change in language learning and processing" (2006: 14). These arguments imply that the impact of the L1 on learning the first foreign language (L2) is fundamentally, i.e. qualitatively, different from the influence exerted by previously learnt languages (i.e. L1, L2 and potentially Ln) on the process of learning a subsequent language. A number of linguistic and psycholinguistic studies support these claims by providing evidence for the existence of qualitative and quantitative differences in processing the third language as compared to the first or second language (Cenoz and Jessner 2000, Cenoz et al. 2001, Hufeisen and Lindemann 1997). Additionally, from a theoretical linguistic perspective, Flynn et al. (2004) argue that the study of L3 acquisition can offer new insights into the process of language learning that neither investigations of the first language nor the second can provide.

From the point of view of speech acquisition, it should be emphasized that multilingual learners have at their disposal a broadened phonetic repertoire, a raised level of metalinguistic awareness and enhanced perceptual sensitivity, which may facilitate the learning of a subsequent phonological system (cf. Gut 2010, Wrembel 2012). In a recent volume on "Universal or diverse paths to English phonology" (Gut et al. 2015), an attempt was made to draw some comparisons between the acquisition of phonology from the SLA vs. TLA perspective. On the one hand, the acquisition of L2 speech has a well grounded tradition of research going back to the 1960s, which initially focused on language learning strategies, to soon emerge as a theoretically oriented field with a specific methodology. As suggested by Gut et al. (2015: 1), current models of L2 phonology (e.g. Best 1995, Flege 1995, Major 2001) are mainly concerned with the relationship between the first (L1) and second language (L2) of the speaker, the role of language universals as well as the influence of non-linguistic factors on the rate, process and outcome of phonological acquisition. On the other hand, research into third language phonology is a very young discipline that has its roots in the late 20th and early 21st centuries. The major difference is that L3/Ln learners have already acquired a foreign language (L2), thus they can resort to previous conscious linguistic knowledge as well as language-learning experience and strategies (cf. Cenoz and Jessner 2000, De Angelis 2007).

Juxtaposing the two acquisition perspectives, Gut et al. (2015) point out that while SLA focuses mostly on the process of acquisition, the focal points in TLA constitute primarily the outcome of this process and its influencing factors. As far as methodology is concerned, the SLA tradition relies on cross-sectional experimental studies, whereas early studies in L3 phonology were mainly longitudinal case studies of multilinguals and only recently has the field turned to the corpus-based, experimental research paradigm. Data elicitation procedures tend to differ as well, since SLA relies mostly on L2 data; however, in TLA it has become a standard to elicit data in all the languages spoken by the subjects. The recent volume edited by Gut et al. (2015), constitutes a valuable endeavour to bridge the theoretical and methodological gap between these two disciplines as well as research on new varieties of English.

Recapitulating, it appears necessary to verify the existing theoretical models and assumptions relevant for second language acquisition and to transform them into a more complex context of learning an additional foreign language. Further, more focal attention should be directed to the acquisition of phonology in a multilingual context due to its being a particularly understudied domain, as pointed out by several scholars (cf. Cabrelli Amaro 2013, Hammarberg 1997).

1.2.1. Terminological debates

Terminological problems and inconsistencies are quite characteristic for an emerging field and it has also been the case for TLA. To start with, the terms 'bilingualism' and 'multilingualism' have been used interchangeably throughout the literature under a widespread assumption that they both refer to the knowledge of two or more languages. Bilingualism has been traditionally used as an umbrella term for any language learning beyond the first language, as evidenced in the literature. For instance, Singh and Carroll (1979: 51) claim that "there is, a priori, no reason to assume that L3 learning is any different from L2 learning. Learning a third language is […] learning just another second language". Grosjean (1992: 51) provided a definition according to which "bilingualism is the regular use of two (or more) languages, and bilinguals are those people who need and use two (or more) languages in their everyday lives". Along similar lines, Sharwood Smith (1994: 7) maintained that the term second language stands for any language other than the first language learnt "a) irrespective of the type of learning environment and b) irrespective of the number of other non-native languages possessed by the learner". Such a 'no-difference assumption' may be ascribed to a general tendency to overgeneralize the notion of a 'second' language and to use it as a cover term (cf. Hufeisen 2000).

Stemming from recent research developments, multilingualism comes to the fore as the term suggested to be used to denote the acquisition of more than two languages and the product thereof (cf. Herdina and Jessner 2002, Jessner 2006). This new approach emphasizes that the distinction is fairly obvious based on the actual meaning of the prefixes 'bi-' and 'multi-' (cf. Cenoz et al. 2003, De Angelis 2007), and the use of the term multilingualism is strongly preferred to bilingualism. Nonetheless, a reversed tendency has also been observed in which the former term is applied as an overgeneralization to refer also to the bilingual context, and bilingualism is treated as a variant of multilingualism. Hoffmann (2001) argues that this is a deliberate practice to emphasize widespread sociolinguistic variety; however, it may generate unnecessary confusion. All in all, a definition of multilingualism provided by the European Commission defines this phenomenon as "the ability of societies, institutions, groups and individuals to engage, on a regular basis, with more than one language in their day-to-day lives" (European Commission 2007: 6), whereas plurilingualism is used to refer to the diversification in the use of languages by an individual (Rothman et al. 2012: 374).

De Angelis (2007) presents four labels that have been coined to refer to the new field of acquisition beyond the second language, namely: Multiple Language Acquisition, Multilingual Acquisition, Third Language Acquisition, and Third or Additional Language Acquisition, and discusses their respective merits and flaws. The first term, Multiple Language Acquisition is problematic as it seems to denote a specific type of concurrent (i.e. simultaneous) acquisition at the exclusion of a more typical sequential setting, in which different languages are learnt not necessarily at the same time. Multilingual Acquisition, on the other hand, is a rather vague term, with 'multilingual' defining the learner rather than the process of acquisition itself. The third term follows the tradition established by earlier acquisition settings of First or Second Language Acquisition; however, according to De Angelis (2007: 11), this notion is too exclusive and does not necessarily imply other languages. Consequently, the author promotes the term Third or Additional Language Acquisition as a more appropriate one, yet it has its limitations related to its length and practical consideration. To the best of my knowledge, Third Language Acquisition emerges as the most preferred option in the literature of the field since it is the actual usage of the term that testifies its validity.

The parameters determining what constitutes a third language have not been universally defined even among researchers dealing with this field. The term 'third language' or L3 has been used variably throughout the literature, either in a chronological sense or based on language dominance. On the one hand, it is common to refer to the speaker's languages in a chronological order of acquisition, i.e. as L1, L2, L3, L4, Ln. Such linear ordering may look appealing for practical reasons, but it raises many problems related to the complex nature of multilingualism, e.g. how to order languages acquired simultaneously? How to account for intermittent or alternating acquisition? Whether to count languages with a very limited command or knowledge limited to a particular aspect? (cf. Hammarberg 2009: 4). A possible solution to this problem is to apply a three-category-distinction, according to which a distinction is made only between L1, L2 and L3. In line with this suggestion, L3 is used as a cover term for any language beyond the first foreign language (L2) since the acquisition of the third language is viewed to be qualitatively different from the L2 acquisition due to the prior experience with a non-native language. It is maintained that any additional languages do not make a radical difference in this process (cf. Hufeisen 1998: 17).

From a dominance-based perspective, the numbering of languages from a multilingual speaker's repertoire is related to their actual proficiency level and frequency of use. According to this view, the ordering can be subject to change, reflecting the dynamic nature of multilingual acquisition and use. However, such an approach seems very vague and rests mostly on rather subjective measures of language dominance.

To reconcile these diverse approaches, a different definition of the third language was put forward, in accordance with which L3 can be used to refer to "a non-native language which is currently being used or acquired in a situation where the person already has knowledge of one or more L2s besides one or more L1s" (Hammarberg 2009: 6). Along these lines, the notion of the L3 does not necessarily correspond to a third language in the order of acquisition, and previously acquired foreign language(s) are subsumed under the category of background languages or L2(s). Further distinctions into L4, etc are not deemed to be justifiable. Another proposal to escape terminological problems stemming from a conventional use of the terms 'first', 'second' and 'third' suggests to replace them with the labels of 'primary', 'secondary' and 'tertiary' languages respectively, to express a more cognitively-based hierarchy between the linguistic systems of a multilingual user in a particular situation (Hammarberg 2009: 7).

1.2.2. Key research areas in TLA

In spite of being a young discipline, Third Language Acquisition has already generated a substantial body of research on a wide variety of topics. The following research areas have been identified as crucial in the field:

multilingual advantage over monolinguals, cross-linguistic influence, early trilingualism, and multilingual education (cf. Jessner 2006).

1.2.2.1. Additive effect of bilingualism

The first of these research areas concerns the alleged linguistic and cognitive advantage that bilinguals and multilinguals enjoy over monolinguals. There is a lot of anectodal evidence concerning additional language learning since many multilinguals report a facilitative effect of previously learnt languages on subsequently acquired systems and an increased ease of acquisition (cf. Cenoz 2003). In one of the early studies, Lambert et al. (1973) demonstrated that multilinguals exhibit enhanced cognitive skills, such as empathy and creative thinking, as compared to their monolingual counterparts. As pointed out by Jessner (2006: 27) "[c]onsiderable evidence shows that the development of competence in two languages can result in higher levels of metalinguistic awareness, creativity or divergent thinking, communicative sensitivity and the facilitation of additional language acquisition".

A critical overview of studies on the impact of bilingualism on the development of cognitive and communicative skills as well as metalinguistic awareness was undertaken by Cenoz (2003). However, due to methodological problems and a range of designs focused on various proficiency levels, the author found it difficult to present a valid comparison of the findings. In conclusion, Cenoz (2003) indicated that the majority of studies on general proficiency tend to exhibit positive effects of bilingualism on additional language acquisition, particularly in the case of typologically related languages. One of the often quoted examples is Ringbom's (1987) study on Finns acquiring English as their L2 or L3, which demonstrated that bilingual children outperformed their monolingual peers. Similar results were reported by Cenoz and Valencia (1994), Lasagabaster (1997) and others studying primarily the acquisition of English as a third language in the bilingual context of Basque-Spanish or Catalan-Spanish, who provided evidence for the additive effects of bilingualism. Nonetheless, mixed results were also observed in the overview of studies on specific aspects of language proficiency which demonstrated less consistency in the findings. For instance, Klein (1995), in his comparison of grammaticality between monolingual and multilingual learners of English, pointed to the advantage of the latter in the rate of learning but not in the outcome. Some studies failed to demonstrate any significant difference between monolinguals and bilinguals acquiring a further language (e.g. Sanders and Meijers 1995). On the other hand, multilinguals have been found to have better communicative competence, lower levels of communicative anxiety and heightened metapragmatic awareness as well as more positive attitudes to language learning (cf. Rothman et al. 2012 for an overview of related studies). Cenoz (2003) concluded that more research is necessary to univocally prove the bilingual superiority in the dynamic and complex context of TLA.

1.2.2.2. Early trilingualism and multilingual education

Early trilingualism appears to be one of the youngest branches of TLA research, which has emerged as a response to a growing number of multilingual families and the resulting acquisition of multiple languages in a naturalistic setting. One of the earliest investigations of this kind included Oksaar's (1978) case study of a child bilingual in Estonian and Swedish acquiring German as an L3, and Hoffmann's (1985) case study of her children – trilingual in German, Spanish and English. Several other studies have appeared since then (cf. Jessner 2006: 26-27), and, interestingly, English seems to enjoy a special status in the development of early trilingualism. While the interest in early simultaneous language acquisition seems to be growing, there is still a lack of research into multiple language acquisition in early childhood (i.e. approximately the ages from 4 to 7), as noted by Franceschini (2009).

A closely related field is that of multilingual education, especially at the primary level. As claimed by Rothman et al. (2012) the growing acquisition of third and additional languages in Europe and North America can be attributed, to a large extent, to current educational policies that started to recognize and promote multilingualism. For instance, on a general level, the Common European Framework of Reference for Languages (CEFR) was created by the Council of Europe to promote individual multilingualism and mobility on the educational, social and economic levels. On a micro-level we have been witnessing a shift in the status of minority languages in several countries which led to the incorporation of these languages in the systems of education, e.g. Basque in the Basque Country, Catalan in Catalunia, Welsh in Wales or instances of Polish in Scottish schools. In an attempt to define what multilingual education actually means, Cummins (2008) proposes that this term should refer to the use of two or more languages, as languages of instruction rather than as school subjects. According to Cenoz (2009), the notion of multilingual education should be defined as goal-oriented since it implies teaching more than two languages under the proviso that a given school actually aims at multilingualism and multiliteracy. Cenoz and Genesee (1998) observe that it can take different forms and may encompass various educational contexts. Current education models are diverse, including double immersion programs (e.g. Hebrew/French in Canada), multiple languages of instruction (e.g. Basque and Catalan alongside Spanish) as well as multiple foreign languages which are subjects of schooling (the majority of European schools at primary and secondary levels). Another model of Content and Language Integrated Learning (CLIL) in which selected subjects are taught in a foreign language different from the mainstream language of instruction, has also proved efficient and is gaining ground in Europe, Asia and the Americas. One of the debated issues concerns the optimum age for the introduction of a third language in the context of schooling (cf. Jessner 2006).

A special volume on teaching and learning in multilingual contexts from sociolinguistic and educational perspectives was edited by Otwinowska and De Angelis (2014). This edited collection illustrates common themes and practices in multilingual education and points to possible solutions of problems, offering a broad geographical representation of schooling contexts. Issues related to multilingual education are discussed in relation to the five types of social settings of multilingual acquisition (adopted from Siegel 2003). Firstly, the dominant L2/L3 setting (i.e. the majority language context) refers to the situation in which the language acquired is the dominant language of the community, whereas people learning it are predominantly immigrants. In such a context, education may result in subtractive bi/multilingualism, leading to the attrition of the native minority language due to negative attitudes towards it. Secondly, in the minority L2 setting the speakers of the dominant language learn the minority language. Such a situation is relatively infrequent and generally takes place in a naturalistic rather than classroom context and it usually results in additive bi/multilingualism. Thirdly, the external setting involves learning a foreign language by the speakers of a language dominant in the region (e.g. Croats learning English in Croatia). Fourthly, the coexisting L2 setting denotes the situation when the respective languages are used in similar domains (e.g. English and French in Canada). Finally, the institutional L2 setting (i.e. the official language context) refers to the situation when the L2 is widely used in some domains or institutions, yet it is an additional language for the majority the population (e.g. Spanish in Catalonia). The authors conclude that each of the identified educational contexts entails different problems and challenges for multilingual education.

1.2.2.3. Research on cross-linguistic influence

The research area that has received the biggest coverage so far is cross-linguistic influence (see section 1.3 for a thorough discussion of this notion). This focal interest stems from potential interactions between the co-existing language systems, which appear to be much more complex and dynamic than in the case of Second Language Acquisition. The TLA perspective widens the possibility of a native-to-target language transfer (L1 → L2, L1 → L3) to include also potential influences of the second and additional language systems (L2 → L3, Ln → L3) as well as to allow for multiple directionality of this interaction (L1 ↔ L2, L1 ↔ L3, L2 ↔ L3, L3 ↔ Ln, etc).

Several attempts have been made to identify predictors of CLI. Kellerman (1979) was the first to propose that similarity is a crucial driving force behind this process. In his understanding, it is the learner's perception of the interlingual distance between two languages, i.e. psychotypology, rather than the actual linguistic typology, that determines transferability. A strong supporter of this claim was Odlin (1989), who pointed out that it is the subjective perception of semantic and categorical relations between linguistic structures in two or more languages that determines the likelihood of transfer-related phenomena. Also Ringbom (1986) interpreted his findings in the light of a greater perceived similarity between L2 Swedish and L3 English (rather than Finnish and English) which determined the source of cross-linguistic influence for his L1 Finnish multilingual learners.

A question arises which stage of acquisition may be particularly susceptible to CLI. In an early observation Ringbom (1986: 155) claims that "[i]t is obvious that the lesser the learner knows about the target language (L2), the more he is forced to draw upon any other prior knowledge he possesses" and concludes that CLI "will be more in evidence at the early stages of learning". This claim has been corroborated in some studies which demonstrated a particularly noticeable influence from the second language at early stages of L3 learning (e.g. Hammarberg and Hammarberg 2005, Wrembel 2010).

Amongst other factors generally acknowledged as influential in determining the sources and directions of cross-linguistic influence most scholars enumerate typological similarity, proficiency level in L2 and L3, recency of use or the L2 status (cf. Williams and Hammarberg 1998). The so called 'L2 status' or foreign language effect refers to the observed tendency to activate the previously acquired second language in the process of the subsequent learning of an additional language and was first reported by Meisel (1983). This effect has been corroborated in a number of studies to date, e.g. Ecke and Hall (2000) found evidence that L1 Spanish learners relied more strongly on their L2 English than the L1 when learning L3 German. Similarly, Lindemann (2000) showed that L1 Norwegian learners relied considerably on their L2 English when learning L3 German. On the other hand, Cenoz (2001), in her study on Basque/Spanish bilingual children learning English as their L3, indicated that the L2 status may be overruled by typological distance as a stronger predictor of CLI. A thorough discussion of other factors affecting CLI can be found in section 1.4.

1.2.2.4. Genesis and development of research in multilingualism

Another relevant area of interest concerns the genesis and development of research into multilingualism. Franceschini (2009) identifies major driving forces behind this, pointing to an increased sensitivity towards socio-cultural diversity and the acknowledgement of the fallacy of the assumption of cultural homogeneity and monolingualism. We have witnessed a shift from a traditional pejorative perception of minority languages or one-language focused studies in isolation. Assuming this new perspective has brought about a much more positive approach towards language diversity and embracing the complex linguistic landscape, both from a vertical (i.e. historical) and horizontal (i.e. territorial) perspective, as noted by Franceschini (2009). Further fields of current academic interest include receptive multilingualism or intercomprehension, based on the mutual intelligibility of languages usually in neighbouring geographical regions. The idea behind fostering receptive multilingualism is that individuals are capable of understanding several languages receptively, although the ability to activate this knowledge in production may be limited.

Moreover, the author identifies also some areas within multilingualism that remain unexplored, including the historical roots of multilingualism, the effectiveness of early acquisition programmes promoted in primary schools, or incidental, unfocused acquisition through mere exposure. Franceschini (2009) points to the dearth of comparative statistical studies of multilingualism aimed at creating dynamic 'maps or indexes of multilingualism' monitoring its development over time and space. The results of such statistical analyses may generate patterns of multilingual practices within families with and without a migration background, profession-related multilingual use, the interdependence between multilingual use and social stratification, etc.

1.2.3. Complexity of TLA

Several scholars have pointed to an inherent characteristic feature of TLA that distinguishes it from SLA, namely, its increased complexity. In the study of SLA we are already faced with numerous factors stemming from sociolinguistic, psycholinguistic and educational perspectives. The acquisition of additional foreign languages further adds to this complexity, as postulated by Cenoz and Genesee (1998), Herdina and Jessner (2002), and Jessner (2006). This complex nature of Third Language Acquisition results from diverse routes of acquisition and learning contexts, a multitude of individual factors related to multilingual learners, as well as the dynamic nature of the process of multiple acquisition.

As far as the routes of acquisition are concerned, several complex patterns emerge upon adding another foreign language to the SLA perspective. As observed by Cenoz (2000) and Jessner (2006) these diverse routes can take the following forms:

1. Three languages can be acquired simultaneously (L1, L2, L3)
2. Three languages can be acquired consecutively (L1 < L2 < L3)
3. Two languages can be learnt simultaneously, following L1 acquisition (L1 < L2, L3)
4. Two languages can be acquired simultaneously, and followed by L3 learning (L1, L2 < L3).

Moreover, the process of acquisition can be subject to interruption and a subsequent restarting, thus resulting in an even more complex pattern (e.g. L1 < L2 < L3 < L2). Further, the context of acquisition can be naturalistic, formal or a combination of both, which when applied to the TLA perspective creates a variety of possible combinations.

An array of individual factors has been identified in the research on Second Language Acquisition including maturational factors, language aptitude, motivation and attitudes, language proficiency, language learning strategies, field dependence or independence, as well as other psycholinguistic and personality factors (cf. Gardner et al. 1997). This repertoire becomes even broader and more complicated if we allow for an interplay of three or more language systems in individual users, i.e. the number of variables related to language status, sociobiographical and psychoaffective factors increases in parallel to the growing number of languages used in particular situations. For instance, the sequence of acquisition will not necessarily correspond to language dominance or proficiency in the particular languages of a multilingual user. The attitudes and motivation related to learning separate languages may differ significantly for individual speakers. Individual variation will need to allow also for the amount of exposure and/or formal instruction in particular languages as well as the context, recency or intensity of their active and passive use. Moreover, a factor that has received special recognition in the multiple acquisition is that of metalinguistic awareness, whose levels are reported to be enhanced in multilingual learners (see section 6.1 for a detailed discussion). According to Cenoz (2000), the number of sociolinguistic factors at play constitutes one of the major differences between second and third or additional language acquisition. Therefore, it appears especially important to create a detailed profile of multilingual users based on their language learning history to be able to account for individual differences in language performance and to point to the independent variables that condition it.

Finally, the complexity of multiple language acquisition stems also from the dynamic nature of the process, characterized by a complex interplay between the existing language systems of a multilingual person. These processes may include, among others, the negative effects of an intense contact between language systems such as language deterioration and attrition or facilitatory processes of re-learning, all of which occur much more frequently as the result of the cross-linguistic influence. The dynamic nature of multilingual acquisition is thus reflected in the non-linearity of the process and the changing roles and status of particular language systems of multilingual users. This has led to the understanding of languages being in a constant flux rather than in a stable state, with subsequent acquisition and attrition viewed as inherent components in the process of language development.

A special nature of the process of multilingual acquisition was also reflected in Cook's (1992) proposal of the term multicompetence, which treats the languages of a multilingual user as part of a larger system rather than as separate entities. According to Cook (2003: 2), "[s]ince the first language and the other language or languages are in the same mind, they must form a language supersystem at some level other than be completely isolated systems". The introduction of the concept of multicompetence contributed also to a broadened understanding of there being a multidirectionality of influence between language systems as well as a more dynamic perspective on language development, e.g. a change in one system may affect other systems as well.

1.2.4. L3 processing

Faced with the results of linguistic and psycholinguistic studies that demonstrate considerable differences in the processing of the third as compared to the first or second language, scholars look for neurolinguistic evidence to support these findings. In an overview of research conducted from a neurolinguistic perspective, de Bot and Jaensch (2013) investigate the assumptions that tend to differentiate between monolingual, bilingual and multilingual processing.

Marian et al. (2012) investigated multilingual processing with the application of the colour-word Stroop test adapted for multilingual participants. They reported the Stroop effect in all three tested languages, with error and disfluency rates being dependent on the level of proficiency. The authors concluded that "[m]ultilinguals were faster and more accurate in the within-language-competition condition than in the between-language-competition condition, indicating that additional processing costs are required when stimulus and response languages differ" (Marian et al. 2012: 2). The findings were interpreted as indicating some quantitative differences in the processing, but there was no evidence that the L3 poses specific qualitatively different demands.

Another source that can inform the theories on the storage of multiple languages in multilinguals' brains includes data from multilingual aphasia. Paradis (1977) distinguished different types of aphasia based on the specific patterns of decline and recovery, including synergistic, differential, antagonistic, successive and selective forms, all of which have been found to depend on a multitude of factors. A comparative analysis of cases of aphasia among bilingual and multilingual patients conducted by Huibregste and de Bot 2002) demonstrated that synergistic, i.e. parallel recovery was the most common and that differential recovery could be attributed mainly to different proficiency levels. The reported patterns of impairment and recovery have not so far confirmed any significant differences between bilingual and multilingual aphasia, but the data are too limited to draw any solid conclusions (cf. Paradis 2004).

Very few studies using neuroimaging techniques have been conducted to date on multilinguals to explore whether the same or different areas of the brain are activated during multiple language use. In Vingerhoets et al.'s (2003) study, the Dutch/French/English trilinguals performed three language processing tasks, including picture naming, verbal fluency and comprehension reading, while whole-head functional magnetic resonance imaging (fMRI) was applied. The results demonstrated that generally the same cerebral areas of the brain were activated during the task performance, although later acquired languages required more neural substrates. The overview of related studies presented by de Bot and Jaensch (2013) implies that largely the same areas of the brain subserve different languages from the multilinguals' repertoire, yet some studies found that the language sites in the brain might be both shared and specific. Furthermore, more brain tissue is necessary for the processing of languages with lower levels of proficiency and automatization. This led the authors to conclude that there is not sufficient neurolinguistics evidence to suggest that the third language is processed differently from the second language and that the neural mechanisms involved are qualitatively divergent. Nevertheless, quantitative difference have been reported, e.g. trilinguals are slower when performing certain tasks and demonstrate cross-linguistic interference reflecting patterns of interaction between all three languages. All in all, based on the literature overview, de Bot and Jaensch (2013: 12) state that "trilingualism provides a specific window on human language processing and provides data that cannot be gathered with bilinguals or monolinguals".

Further studies may also be necessary to investigate the role of attention control (i.e. attentional flexibility) and inhibition (i.e. controlled suppression mechanism) in phonological processing in third language acquisition, in parallel to the ones conducted from the second language acquisition perspective (e.g. Darcy et al. 2014). Their findings demonstrated that more efficient attention control and inhibitory skills result in the enhanced processing of phonologically relevant acoustic information in the second language input, and consequently may contribute to greater accuracy in foreign language speech perception and production.

1.3. Cross-linguistic influence

The present section will focus on the phenomenon of cross-linguistic influence (CLI), which seems crucial in the research on TLA, by outlining its specificity, typology and sources. The term cross-linguistic influence (CLI) was coined by Sharwood-Smith (1983) and Kellerman (1984) as a cover term to refer to transfer-related phenomena in a more theory-neutral manner. CLI was intended to include a broader range of linguistic influences stemming from language contact, such as transfer, interference, borrowings as well as language attrition or loss caused by the coexistence of several language systems.

From the Second Language Acquisition perspective, language transfer has been traditionally viewed as a one-to-one type between the native and the target language. With the development of research into third or additional language acquisition, a broadened understanding of this phenomenon has been posited advocating a simultaneous influence of more than one previously acquired languages. The earliest mentioning of such broadened conceptualization of transfer encompassing other non-native languages apart from the native tongue can be traced back to the 1980s. To illustrate this, Gass and Selinker (1983: 372) view language transfer as "the use of native language (or other language) knowledge […] in the acquisition of a second (or additional) language". Similarly, Sharwood-Smith (1994: 198) defines it as "the influence of the mother tongue on the learner's performance in and/or development of a given target language; by extension, it also means the influence of any 'other language' known to the learner on that target language".

Such a broadened understanding of transfer, extending beyond L1 influence, has been embraced by scholars investigating third or additional language acquisition (e.g. Cenoz et al. 2001, De Angelis 2007). This resulted in a new conceptualization of transfer-related phenomena, encompassing different interactions between non-native languages and a simultaneous influence of more than one language on the target language currently being acquired. The results of empirical investigations carried out from the multilingual perspective have challenged well-established assumptions that used to identify the native language as the only or the prevailing source of influence and led to modification of the existing theoretical models. Consequently, the traditional notion of one-to-one transfer associated with SLA was substituted with a many-to-one type or the so called 'combined cross-linguistic influence', as suggested by De Angelis (2007: 21) (see section 2.3.3). The fallacy of the term transfer, as applied specifically to phonological acquisition of the second and third language, is discussed by Dziubalska-Kołaczyk and Wrembel (forthcoming).

The following section will discuss the phenomenon of cross-linguistic influence, focal for the present contribution, in terms of its typology as well as sources and directionality.

1.3.1. CLI typology

Trying to account for the complexity of cross-linguistic influence, Jarvis and Pavlenko (2007) put forward a proposal of a detailed categorization based on ten various dimensions: (1) the area of language knowledge (e.g., phonological, semantic, lexical transfer, etc.), (2) directionality (forward, reverse, lateral, multidirectional transfer), (3) cognitive level (linguistic vs. conceptual transfer), (4) type of knowledge (implicit vs. explicit), (5) intentionality (intentional vs. unintentional transfer), (6) mode (productive vs. receptive), (7) channel (aural vs. visual), (8) form (verbal vs. nonverbal), (9) manifestation (overt vs. covert) and (10) outcome (positive vs. negative) (2007: 20). These dimensions will be discussed in more detail in following paragraphs.

As far as the area of language knowledge is concerned, the present investigation is focused on phonological transfer, which has, so far, been dominated in TLA research by other fields, particularly morphosyntax and lexis.

One dimension that needs special elucidation is that of directionality. The distinction between 'forward transfer' as well as 'reverse' or 'backward transfer' has been used rather conventionally in the SLA literature (e.g., Gass and Selinker 2001) to refer mostly to the transfer from the native to the target language L1 → L2 (i.e. forward) and from the target to the native language: L2 → L1 (i.e. reverse). In the case of Third Language Acquisition, a simple extension of these terms to refer to, for instance, second to third (L2 → L3) or fourth to third language (L4 → L3) would not be fully adequate due to the commonly non-sequential order of acquisition and other factors influencing the actual status or dominance of languages such as proficiency or recency of use. Therefore, a complementary term of 'lateral transfer' was introduced by Jarvis and Pavlenko (2007) to refer to any influence of a non-native (or post-L1) language on another non-native language (e.g. L2 → L3, L3 → L4). Moreover, 'bidirectional or multidirectional transfer' was proposed to cover cases when linguistic systems from the multilinguals' repertoire function simultaneously as source and recipient languages, i.e. synchronous forward and reverse transfer (L1 ↔ L2) or synchronous bidirectional lateral transfer (L2 ↔L3) (Jarvis and Pavlenko 2007: 22).

With respect to the cognitive level, a traditional view of transfer entailed transferring mental representations from one linguistic system to another. A more modern stance questions this assumption, juxtaposing a view in which transfer contributes to the formation of mental links between two or more languages. A commonly encompassed scenario envisages that linguistic systems of a multilingual user can influence one another at a number of cognitive levels. Consequently, CLI is postulated to occur at the level of conceptual representations, semantic representations or linguistic representations either selectively or simultaneously (Jarvis and Pavlenko 2007: 23).

Moreover, Jarvis and Pavlenko (2007) are convinced that the distinction between implicit and explicit linguistic knowledge, corresponding to how languages are stored and processed in the mind (cf. Paradis 2004), may have also significant implications for cross-linguistic influence. Accordingly, they propose to identify types of CLI related to implicit and explicit knowledge with the proviso that more research with the application of converging methodologies is necessary to explore this issue further.

Another distinction in the CLI categorization is made based on the intentionality of transfer. On the one hand, CLI may function as an intentional communicative strategy resulting from formed mental associations between languages or interlingual identifications. Yet, on the other hand, CLI may also take the form of unintentional language switches.

So far, research into the intentionality of cross-linguistic influence in TLA has been quite limited, with the exception of Hammarberg (2001) and Williams and Hammarberg (1998), who identified the multilingual subject's intentional strategy to avoid L1 transfer in third language production resulting in a reliance on the second language.

Mode is yet another CLI dimension, referring to the division into four language skills; i.e. productive, including speaking and writing, vs. receptive, involving listening and reading. This basic distinction is seen as having implications for types of cross-linguistic transfer as it may concern either speech production or comprehension. Apart from modality, various channels of communication affect the typology of CLI, ranging from aural to visual transfer. Along similar lines, the form dimension of CLI refers to the distinction between verbal vs. nonverbal linguistic performance. With a natural prevalence for transfer in the verbal form, recent research interest has been shifted to include also bimodal or gestural communication as an indicator of foreign language competence (e.g. Gullberg 2006).

Cross-linguistic influence has been also postulated to have different manifestations, including overt vs. covert types. The former are understood as interlingual identifications made by the language users between forms and meanings in two or more languages, whereas the latter refer to instances in which such overt identifications have not been made but rather they result from avoidance strategies (cf. Ringbom 1987 for a more thorough discussion thereof). Finally, an outcome-related distinction into positive vs. negative transfer is postulated. The use of these terms can be traced back to Selinker (1969) and they have been applied widely in many SLA debates. Traditionally, focal attention has been directed to negative transfer due to a preoccupation with errors, whereas from a more recent perspective scholars have started to account for more general effects of CLI irrespective of the outcome (e.g. Cook 2002).

The phenomenon of cross-linguistic influence has started to be investigated extensively in the context of multilingual language acquisition. For instance, an edited volume by Gabryś-Barker (2012) discusses various aspects related to the syntactic and lexical development of foreign language learners from diverse L1 backgrounds, focusing on the areas that are particularly prone to influence induced by multiple language contact.

1.3.2. Factors, sources and directionality of CLI

Having discussed the typology of cross-linguistic influence, I would like now to elaborate more on its directionality and possible sources, including transfer from the native language (L1 effect), from other non-native languages as well as combined CLI.

In the SLA tradition, the first language is considered as a strong source of phonological interference due to the established neuro-motor routines. Theoretical accounts of this phenomenon rely primarily on the transfer of phonetic features characteristic of the first language to the learner's second language, resulting in foreign accented speech. Models of L2 speech perception and production have identified the effect of the so called L1 filter on the learner's performance in the second language. For instance, the Perceptual Assimilation Model (PAM) by Best (1995) and its extension PAM-L2 by Best and Tyler (2007), aimed to explain how monolinguals and L2 learners perceive and assimilate non-native sounds in terms of articulatory gestures. One of the major tenets of this model is that non-native sounds are perceived through a comparison with L1 sounds which are closely located in the phonological space, and categorized on the basis of goodness of fit as exemplars of a native category (Best 1995). In another influential model, i.e. the Speech Learning Model (SLM), Flege (1995) postulates that second language acquisition applies mechanisms and processes used in first language acquisition as they remain intact over the life span. Flege further elaborates that the category formation for second language sounds may be blocked by the process of 'equivalence classification', in which perceptually similar segments in the L1 and L2 are assimilated to a single phonetic category. Consequently, the L1 filter may result in inaccurate perception and accented production in the second language.

Early approaches to L3 phonology have partially embraced this stance of L1-dominated transfer, while a wider recognition of non-native languages as sources of influence has come with the development of the field. Quite a convincing explanatory account was provided by Hammarberg and Hammarberg (2005: 17), according to which "[…] influence from L1 on articulatory settings is a basic constraint on articulation which tends to be persistent in language learning, whereas the reliance on L2 settings is a coping strategy which the learner resorts to at an initial stage when the phonetic form of L3 is too unfamiliar to master".

Other arguments postulating the possibility of cross-linguistic influence from non-native sources are based on psycholinguistic grounds and allude to simultaneous phonological activation of all the languages in the mind of the multilingual user. Following this line of argumentation, phonetic elements from other foreign languages from the multilingual's repertoire may unintentionally appear during speech production in spite of the inhibitory control performed by the articulator (cf. de Bot 1992, Poulisse and Bongaerts 1994). The reduced inhibition of foreign elements may occur due to the perception of similarity between particular languages (e.g. L2 and L3), a lower proficiency level or the impact of psycholinguistic factors of language anxiety or tiredness on the part of the multilingual learner.

The phenomenon of activating another non-native language, which is usually the L2, in the production of the third language, has started to be referred to as the 'foreign language effect' or the L2 status, and it has attracted a lot of scholarly attention (e.g. Cenoz 2001, Fernandes-Boëchat and Siebeneicher Brito 2008, Hammarberg and Hammarberg 1993, 2005, Jessner 2006, Williams and Hammarberg 1998). Trying to account for this source of cross-linguistic influence, De Angelis and Selinker (2001: 56) claim that there is some kind of 'a potential foreign talk' or 'foreign language cognitive mode' that facilitates the path of interlanguage transfer.

The phenomenon of cognitive associations between foreign languages has been reported quite extensively in the literature. For instance, Cohen (1995), in his analysis of case studies of trilingual speakers, observed that "the learner's mind would go into a 'foreign language' mode in what would appear the dominant foreign language rather than the target one" (Cohen 1995: 102). In turn, according to De Angelis (2005), two interacting constraints concur in blocking the influence of the native language, including the perception of correctness and the association of foreignness. The first constraint predicts that the transfer from L1 into the third language will be blocked because of the multilingual learner's realization of the incorrectness of L1 forms and an increased acceptance for non-native forms in the target language. On the other hand, the second constraint, i.e. the association of foreignness, assumes that non-native languages will be assigned a common status of 'foreign languages' and consequently, the use of non-native forms in the L3 will be favoured through cognitive associations over the native forms (de Angelis 2007: 29).

Numerous hypotheses have been put forward in the literature to explain this switch to a foreign language mode. Firstly, it is indicated that the mechanisms responsible for the acquisition of the first language are inherently different from those characteristic for foreign language learning. The process of learning a second or a third language usually takes a similar route of acquisition (i.e. formal instruction), which differs from the naturalistic acquisition of the mother tongue. As a result, it is suggested that L3 phonological learning may lead to the reactivation of the L2 acquisition mechanisms (cf. Hammarberg 2009). Secondly, the foreign language effect tends to be interpreted as a coping strategy that is resorted to at the initial stages of the phonological acquisition of the third language when the L3 phonetic form is too unfamiliar for the learner. Along these lines, Hammarberg and Hammarberg (2005) maintain that the coping strategy is able to override temporarily the basic constraint of L1 transfer, yet it later diminishes with the development in L3 proficiency. Thirdly, the predominant role of the second language as the external supplier in the process of L3 learning may be also accounted for in terms of psycho-affective variables. Literature reports indicate learners' attitudes as influencing factors reflected, for instance, in the desire to suppress the L1 as 'non-foreign' and to apply a 'foreign language strategy' instead, as observed by Hammarberg (2009). Consequently, the interplay of two processes of suppressing the L1, on the one hand, and activating the L2, on the other, results in a strong reliance on the second language as a significant source of cross-linguistic influence. Other explanations for the prevalence of the second language as the source of CLI are provided by research on the multilingual mental lexicon. In this framework, Hall and Ecke (2003) propose that there is a stronger transfer between the third and second languages, rather than between the third and the first, which can be attributed to stronger links between foreign languages in the speaker's mind.

In his account of third language learning, Pyun (2005) proposed four different sources of knowledge: (1) native language (L1) rules; (2) L2 rules; (3) target language (L3) rules; and (4) 'interrules', which bridge the gap between the previously acquired languages and the ones being currently learnt. This laid the foundation for the proposal of a combined cross-linguistic influence, which allows for multiple sources of transfer, including any previously acquired language systems (cf. De Angelis 2007, see sections 2.3.3 and 3.2.5).

Apart from focusing on the potential sources of cross-linguistic influence, the scholarly discussion concerns also its effects (i.e. negative vs. positive transfer) as well as directionality. The traditional conceptualization of cross-linguistic influence as a mostly restricting factor related to negative transfer and interference has been subject to major changes. A recent trend in foreign language acquisition has emphasized the beneficial effects of positive transfer reflected e.g. in the successful suppression of negative transfer from the L1 to the L3 or a correct application of phonetic phenomena transferred from another foreign language (L2) to the currently acquired one (L3). Marx and Mehlhorn (2010), in their practical discussion on encouraging positive phonological transfer from L2 to L3, plead to make use of declarative knowledge from various linguistic sources and to incorporate the procedural knowledge developed through the learning experience with previously acquired foreign languages.

The major trend in research focuses on progressive transfer, i.e. from an earlier to the subsequently acquired language, e.g. L1 → L2, L2 → L3, with fewer studies acknowledging the possibility of a regressive transfer, e.g. L2 → L1, L3 → L2 (e.g. Cabrelli Amaro 2013, Sypiańska 2013, Wrembel 2011a, see section 3.2.6 for a more detailed discussion). However, the research perspective has been considerably widened through the recognition of both the multiple sources of transfer as well as the multidirectionality of cross-linguistic influence. Further, the nature of cross-linguistic influence has been found to be conditioned by an array of factors, which will be outlined in more detail in the following section.

1.4. Factors affecting third language acquisition

Research on third language acquisition conducted to date has identified a number of factors that contribute to predicting the source, direction and relative strength of the influence of previously learnt languages on the subsequently acquired language systems. Among the factors that come to the fore in L3-oriented literature discussions are language distance, psychotypology, target/source language proficiency, the sequence of acquisition of particular languages, recency of use, length of residence and metalinguistic awareness (cf. De Angelis 2007, Cenoz 2001). The following subsections will focus on selected variables that are of particular relevance for the present contribution.

1.4.1. Language distance and psychotypology

With regard to the factor of language distance or, in other words, typological proximity, scholars generally agree that cross-linguistic influence is most likely to occur between languages which are closely related rather than those which are not (e.g. Williams and Hammarberg 1998, Cenoz 2001, De Angelis 2007). Examples of pertinent research findings indicate that multilingual speakers tend to be influenced mostly by those languages from their linguistic repertoire that are in the closest proximity to the target language. Nonetheless, there have been also some reports, albeit less frequent, of reliance on more distant languages (cf. De Angelis 2007).

Typological proximity can be viewed in a two-fold manner: on the one hand, as an objective formal measure of a genetic relationship between language families, or, on the other hand, as learners' subjective perception of that language's distance. The notion of perceived language distance, i.e. psychotypology, was proposed by Kellerman (1987). He put forward a claim that language transferability is conditioned by two constraints, namely, psychotypology and prototypicality. According to Kellerman, more prototypical forms and features in the source languages lead to a higher degree of cross-linguistic influence affecting the target language, particularly if these languages are perceived to be similar.

Ringbom (2002), in turn, makes a distinction between three tiers of transfer, involving the overall level, the item level and the system level. The first type refers to the overall perception of similarity between the language systems of a multilingual user and Ringbom is of the opinion that it has a facilitative effect on learning. The second type, i.e. item transfer, is based on the established interlingual identifications and learner's reliance on form rather than meaning, and it may result in a positive or negative transfer in the form of one-to-one mappings. The third type, i.e. system transfer, is related to the identifications of the identity of meanings in cross-linguistic items and may take the form of loan translations or semantic extensions. Along similar lines, De Angelis (2007) points to yet another distinction that was drawn between relatedness and formal similarity. The former stands for a genetic affiliation or typological proximity between languages belonging to the same or related language family or group (e.g. Germanic or Romance languages). Formal similarity, in turn, refers to the identification of a similarity between unrelated languages with respect to some language components or features (e.g. the agglutinative characteristics of Swahili and Finnish identified by learners in spite of the lack of any genetic relationship between these languages, as reported by Ringbom (2003). Recapitulating, language distance, both systemic and perceived, appears to be a complex variable with a powerful explanatory potential related to the source and directionality of cross-linguistic influence in TLA.

1.4.2. Proficiency in target and source languages

Another factor commonly acknowledged in the L3 literature as conditioning cross-linguistic influence is that of the proficiency level both in the target and source languages. Scholars are quite unanimous in their claim that CLI is more likely to take place at the early stages of acquisition when the proficiency level in the target third language is rather low; however they don't preclude it happening at more advanced levels as well.

L3 research results to date have attested several instances of CLI at a low level of advancement in the target language when multilingual learners tend to resort to transfer as a coping strategy (e.g. Hammarberg and Hammarberg 2005, Odlin 1989, Wrembel 2010). As observed by Odlin (1989), the transfer characteristic for the low proficiency level in the target language is usually negative since it is driven by the necessity to fill gaps in knowledge. On the contrary, the positive type of transfer typically occurs at more advanced stages of acquisition when learners benefit from their previous linguistic knowledge to a wider extent. Furthermore, scholars indicate that the proficiency level in the source language may also play a significant role in determining the cross-linguistic influence; however, relatively few empirical studies to date have explored this issue in depth. The conclusions also are mixed, as some authors claim that non-native languages can constitute CLI sources irrespective of how proficient the multilingual learners are in these source languages (e.g. De Angelis 1999, 2005, Rivers 1987). Other scholars maintain that the proficiency threshold level in the source non-native language must be sufficiently high in order to exert an influence on the target foreign language (e.g. Gut 2010, Fernandes-Boëchat 2007, Ringbom 1987). All in all, it appears that the proficiency level in both the source and target languages is a significant factor in shaping the patterns of cross-linguistic influence.

1.4.3. Language use and exposure

Other CLI predictors are related to foreign language use and exposure. These factors involve, more specifically, the recency of language use as well as the length of residence in a given country and exposure to a foreign language environment, which are generally found to influence the amount and type of transfer. To begin with, the recency of use has been identified as one of potential conditioning factors in some studies on multilingualism conducted over time (e.g. Hammarberg and Hammarberg 1993, 2005, Vildomec 1963). The underlying assumption is that recent language use facilitates the occurrence of potential influence due to the previous activation of some linguistic information stored in the mind of a multilingual. Nonetheless, some counterevidence has also been found in reports of transfer from languages that have not been in use for a long time (cf. Rivers 1979, De Angelis 1999). As far as the exposure to the non-native language environment is concerned, it has been maintained that a long period of residence in such an environment could exert an influence on the amount and type of CLI in L3 (cf. Vildomec 1963). De Angelis (2007: 37-38) presented an overview of studies whose findings provide strong support for Vildomec's claim. Another factor that has been widely acknowledged in the third language acquisition literature is related to metalinguistic awareness, but this will be discussed at length in Chapter Six (section 6.1).

1.5. Conclusion

Recapitulating, Third Language Acquisition has emerged as a separate field of inquiry, which is fundamentally different from the related area of Second Language Acquisition. The present chapter aimed to characterize this novel perspective by outlining its genesis and development; discussing major differences between the acquisition of the first vs. subsequent foreign languages; explaining the rationale behind the inherent complexity of TLA and presenting its key research areas. Finally, the phenomenon of cross-linguistic influence, focal for this dissertation, was elaborated on, allowing for the CLI typology, sources, and influencing factors. As an independently developing scholarly field, third language acquisition requires new approaches to its theoretical conceptualization and modelling, which will be presented in the following chapter.

Chapter 2

Models of multilingual acquisition

2.1. Introduction

Third language acquisition is a young field of studies, whose theoretical foundations are being shaped by related disciplines such as bilingualism, second language acquisition or psycholinguistics. It has recently gained recognition as a discipline in its own right; however, it appears to be still in a transitory state, which has implications for the undertaken attempts at modelling.

To present this status quo, I decided to embrace a framework illustrating the phases of transfer research proposed by Jarvis and Pavlenko (2007: 5-6), as pertaining to the development of any phenomena in language and cognition. Transfer research has been characterized as progressing through four major phases and I intend to expand it to third language acquisition research more generally. According to the framework, Phase 1 involves the recognition and investigation of a given phenomenon as a potential explanation or a factor (independent variable) that affects other processes (e.g. SLA). Primary research concerns, at this stage, include identifying the phenomenon, defining its scope and quantifying its effects. At Phase 2, the phenomenon begins to be investigated as a primary process (i.e. as an explanandum or dependent variable) and the main research areas concern e.g. identification of its causes and constraints, or verification of its effects. At Phase 3, theories designed to explain the phenomenon are developed and embedded in social, situational or cognitive frameworks. Primary research concerns at this stage feature the development of theoretical models and empirical testing of specific hypotheses. The final stage, Phase 4, implies a comprehensive understanding of the phenomenon, and it involves complex neurophysiological accounts of its functioning. Trying to position third language acquisition research within this framework, we would need to point to a transition period between Phases 2 and 3, with a lot of overlap between them. As evidenced in the literature, this strand of research has gained recognition in the late 20th and early 21st centuries and started to be investigated as a phenomenon in its own right, with key research areas encompassing the identification of sources and directions of cross-linguistic influence, or the verification of multilingual advantage (cf. Chapter 3 for an overview of selected literature). Moreover, the field of third language acquisition has attracted also theoretical interest that gave rise to the emergence of competing theoretical models, explanatory accounts and empirical investigations. This phase has not been completed and the design of theoretical models appears to be still under way, while the grounds for Phase 4 have already been laid with an increasing number of neurolinguistics investigations exploring the complexity of multilingual acquisition.

This chapter aims to discuss the theoretical underpinnings of the present contribution and is composed of four major sections, three of which present different approaches to modelling third language acquisition. Various types of L3 modelling were subdivided by the present author into three categories including: classical models of foreign language processing (section 2.2), tentative models of multilingual speech (section 2.3.) and current theoretical models put forward exclusively for third language acquisition (section 2.4). The final section elaborates on the theoretical conceptualization of the present series of studies by outlining the research hypotheses and predicted outcomes.

2.2. Classical models of foreign language processing

The majority of the models that have been proposed in the realm of multilingual acquisition so far stem from a psycholinguistic perspective and research on bilingualism and Second Language Acquisition. In the present section, classical models of foreign language processing that have been adopted from the related domains and applied to multilingualism shall be outlined, including de Bot's (1992, 2004) multilingual production model, Green's (1986, 1998) activation/inhibition model, Grosjean's (1998, 2001) language mode hypothesis, and Herdina and Jessner's (2002) Dynamic Model of Multilingualism.

2.2.1. De Bot's multilingual production model

The model of bilingual production proposed by de Bot (1992) was developed on the foundations of Levelt's (1989) classical model and then further extended to multilingual acquisition (de Bot 2004). De Bot closely followed Levelt's original model for monolingual speech, according to which language processing occurs in three successive subsystems, namely, the conceptualizer, the formulator and the articulator. The conceptualizer is held responsible for converting communicative intentions into pre-verbal messages, based on its access to extralinguistic knowledge about the world. The message is transmitted to the next tier, i.e. the formulator, which in turn, has access to the lexicon consisting of a lemma subcomponent (i.e. semantic and syntactic information) and a lexeme subcomponent (i.e. forms of words). At the formulator level two separate processes take place, namely grammatical and phonological encoding, in which a surface structure and a phonetic plan are produced and then fed into the next subsystem, i.e. the articulator.

In his extension of Levelt's language processing model to account for bilingual acquisition, de Bot (1992) embraced the idea of three language-specific overlapping subsets and merely added an external language node that performs a monitoring function. This language node is expected to monitor the state of activation of the various languages of a multilingual speaker, as stipulated by Green (1986), and compares the intended language with the actual output. When specifying the criteria that the model should fulfil, de Bot (1992) emphasized that it is to account for instances of cross-linguistic influence, for different levels of proficiency in respective languages and that it should deal with a potentially unlimited number of languages. De Bot proposed some modifications to Levelt's original model. For instance, he challenged the assumption that macroplanning in the conceptualizer is language-specific and suggested that this language-specific information needs to be added to the general pre-verbal message. At the formulator level, his initial proposal of dual scenarios (i.e. a common vs. separate lexicons for two languages) evolved into a solution advocating the common storage of some language elements alongside a separate storage of others, depending on the factors of proficiency levels and linguistic distance. This part of the proposal was subject to several modifications, and thus it appears less consistent. Further, de Bot does not elaborate on his claim that the bilingual production model should be capable of accounting for an unlimited number of languages, and it remains unclear how many parallel speech plans such a system would be able to produce and actually execute. Finally, the idea of separate formulators is not maintained at the articulator level, where the phonological encoding takes place. De Bot is of the opinion that bilinguals share a common store of syllables and he follows Levelt's conceptualization of basic units of speech being syllables rather than sounds. He does not specify, however, how these sounds from various languages are distinguished during the actual speech process and if there are some language-specific labelling mechanisms at play. This model was further adapted to multilingualism, as de Bot (2004) claimed that there is no necessity for developing a separate model since the present understanding of multilingual processing is still rather limited.

Another attempt to apply Levelt's processing model to the multilingual perspective was undertaken by Clyne (2003). In his model of plurilingual processing, he tried to integrate sociolinguistic and psychological aspects to account for specific language choices and speakers' multiple identities. On the whole, Levelt's psycholinguistic model of speech processing turned out to be very influential in shaping theoretical perspectives in multilingual acquisition. It can be argued, however, that this psycholinguistic framework has been tested empirically but mostly based on monolingual data, and thus further research is necessary to validate it further.

2.2.2. Green's activation/inhibition model

In turn, Green based his model of the control of speech (1986) on his research into multilingual aphasia recovery patterns and code-switching. The basic tenet of the model is that languages of a bi-/multilingual person exhibit different levels of activation rather than being in a binary on/off state. These varying activation levels include: (1) the selected level, i.e. a language controlling the actual output at a given moment, exhibiting the highest activation level; (2) active, i.e. a language that takes some part in the ongoing speech processing, e.g. through code-switching; (3) dormant, i.e. a language that does not interact in the current speech process, but is stored in the long-term memory. Therefore, it transpires that multilingual speakers' languages are continually activated but to a different degree. Green's account of how control is executed is based on the assumption that both activation and inhibition operate concurrently during the process of speech production. The state of activation is determined by the frequency of use of particular languages in the sense that frequently used languages can remain active in the ongoing processing in a parallel language, whereas those infrequently used tend not to exert such an influence. This claim seems somewhat controversial as there has been some evidence of cross-linguistic influence also from the so called dormant languages, as attested in the literature (cf. De Angelis 2007).

A later development was Green's (1998) inhibitory control model, in accordance with which there are multiple levels of control that act as inhibitors for potential competitors in the process of production at the lemma level. This monitoring, attentional system supervises the established schemata; however, sufficient resources are necessary in order to control the system. The model can account for asymmetrical switching costs since it takes longer to switch into a dominant language, which needs to be more suppressed. Furthermore, it is indicated that second language use requires more energy to monitor the activation and inhibition processes since foreign language systems are not as automatized as the first language system. Hypothetically, increasing the number of activated languages in the case of a multilingual user would entail the resource generator to produce supplementary resources to be able to exercise additional control. Green is of the opinion that there is a limit to the number of languages that can become activated at a given point in time; however, he fails to define these limits in more specific terms (cf. De Angelis 2007: 73-74). Another weakness of the model is that it is constructed around the notion of intact systems, according to which any deviation from a native-like norm is viewed as an inability to exercise control over the system. Such an assumption appears to be an expression of a monolingual bias in research, a stance that has received a lot of criticism in the recent literature (e.g. De Angelis 2007).

In one of the most recent publications, Green and Wei (2014) developed the model further and elaborated on the control process model of code-switching in bilinguals. The authors propose that the speech planning mechanism is governed by cognitive control processes (CPs) which are suited optimally to particular kinds of code switches. The access to the planning layer is governed by different control modes, which monitor two language gates, including competitive control and cooperative control, the latter of which is subdivided into coupled and open control modes. The coupled mode allows alterations and insertions, while the open control mode is a prerequisite for dense code-switching. In the competitive coordination, on the other hand, one language schema inhibits the other. To account for the serial order articulation, the authors embrace the competitive queuing networks as a neuroanatomically feasible framework. This contribution of Green and Wei explores various predictions of this model of cognitive control processes and offers valuable implications for the research on code-switches in bilinguals.

2.2.3. Grosjean's language mode hypothesis

Another model of bi- and multilingual language processing was put forward by Grosjean (1998). According to his hypothesis of the language mode continuum, a bilingual speaker can choose a base language, i.e. the most highly activated language at a given point in time, as well as the state of activation of other languages and language processing mechanisms. Consequently, at one end of the continuum a person can be in a totally monolingual mode, whereas at the other – in a bilingual or trilingual mode, with various levels of activation of the pertinent languages. The position on the language mode continuum, in turn, is conditioned by the variability of speech situations. Grosjean (2001) acknowledged several psycholinguistic and social factors that influence the language mode, featuring, among others, language proficiency, the speakers' language mixing habits, their socioeconomic status, the degree of formality, the purpose of the interaction, interlocutors, e.g. the presence of monolinguals, etc. In his more recent version of the language mode hypothesis, Grosjean (2001) explicitly extended the model to multilingualism, by claiming that it can be applied to several languages that can be activated to various degrees during the ongoing speech process and, thus, influence the actual language output.

Some scholars argue that their findings provide counterevidence for the language mode hypothesis. For instance, Dijkstra and van Hell (2003) tested the hypothesis on trilingual speakers and found that although the participants were intentionally set in a monolingual mode, they showed evidence of a parallel activation of other languages when processing cognate words as opposed to non-cognates. Grosjean, however, was able to partially account for evidence of this type, maintaining that "bilinguals rarely deactivate the other language totally" (1992: 59), and even if they are in a monolingual mode, they still exhibit some residual activation in the background language/s. On the other hand, Dewaele (2001) interpreted his findings as providing support for Grosjean's language mode hypothesis, since the proportion of mixed utterances used by his multilingual participants differed significantly as a function of the formality of the situation and the different position on the mono vs. bilingual mode continuum. Nonetheless, the hypothesis has not been sufficiently tested from a multilingual perspective and requires further empirical validation, as indicated by De Angelis (2007: 80).

2.2.4. Herdina and Jessner's Dynamic Model of Multilingualism

To account for the process of multilingual acquisition in a more holistic manner, Herdina and Jessner (2002) proposed a Dynamic Model of Multilingualism (DMM), applying the assumptions of dynamic systems theory that has been vibrant in the sciences for decades. The major tenet of DMM is the non-linearity of language growth, a dynamic change over time and an interdependence between language systems. This broadened understanding implies that the acquisition of an additional language results in the development of new qualities in the multilingual system, which is also influenced by socio- and psycholinguistic variation. According to Herdina and Jessner (2002), multilingual proficiency entails a cumulative measure of psycholinguistic systems in contact, the cross-linguistic interaction between them and the so called multilingualism factor (M-factor). To illustrate the concept of multilingual proficiency, the authors of the model put forward the following formula (cf. Jessner 2006: 33):

$LS_1 + LS_2 + LS_3 + LS_n + CLIN + M = MP$
LS: language system
CLIN: cross-linguistic interaction
M: M(ultilingualism)-factor
MP: multilingual proficiency

In this attempt to adopt the dynamic systems theory to multilingual acquisition, Herdina and Jessner (2002) emphasize the change of quality in additional language learning due to the catalytic effect of the third language, and enhanced metalinguistic awareness and metacognitive strategies that all form part of the M-factor. Furthermore, the concept of CLIN is understood as encompassing transfer related phenomena such as code-switching, borrowings, interference and other non-predictable dynamic effects of cross-linguistic interaction between the language systems of a multilingual speaker.

Along these lines, de Bot (2012), in his contribution on rethinking multilingual processing, juxtaposes the main characteristics of current models of multilingual processing with the features of complex dynamic systems. One of the basic assumptions of the dynamic perspective on language processing is that such a complex system as the multilingual mind keeps interacting continually with the environment and, consequently, changes continuously over time. In a detailed overview of dynamic systems characteristics, de Bot (2012: 82-83) emphasizes that dynamic systems are always embedded within other systems and fully interconnected rather than functioning autonomously as part of modular processing, as viewed in the traditional perspective. Further, complexity emerges out of the application of simple procedures, yet the outcome of the development cannot be predicted due to the interplay of interacting variables. The development of the system is fuelled by the interaction with the environment and the principle of internal self-organization. The inherent variability of dynamic systems stems from individual variation, both within and between individuals. De Bot criticized some of the principles of the traditional language processing 'steady state' models, including the modular approach, no internal feedback and feedforward claim, isolated elements as objects of study and the assumption of invariant and abstract representations. He further argued in favour of assuming dynamically based models and outlined their characteristics. According to de Bot (2012), models of multilingual processing should embrace a major tenet that languages do not exist in the human brain as entities, but rather as interconnected networks of subsystems, and that they allow for continuous change over time and variability as their core elements. Moreover, language should be seen as "distributed, situated and embodied; therefore linguistic elements should not be studied in isolation but rather in interaction with the larger units of which they are part and the smaller units they consist of" (2012: 90). In conclusion, de Bot emphasizes that language use appears to be one of the most complex processes of human cognition; therefore it requires an adequate approach to account for its complexity, and dynamic models offer a promising perspective for investigating the development of multiple languages over the course of time.

2.3. Models of multilingual speech

The present section aims to outline some proposals for modelling multilingual speech that have become widely recognized in the third language acquisition literature. They are not based on any psycholinguistic models of speech processing, nor stem from any specific linguistic theory. These models represent a mostly data-driven approach, tend to be less elaborate and assume the form of proposals explaining the observed patterns in third language acquisition. The three models of multilingual speech that were selected for a more detailed presentation include Hufeisen's (1998) Factor Model, Hammarberg's role-function model, and De Angelis' combined CLI.

2.3.1. Hufeisen's Factor Model

The Factor Model was originally introduced by Hufeisen (1998) and further developed by Hufeisen and Marx (2007). Its major assumption is that particular factors are added in the course of the acquisition of subsequent languages, thus leading to an increased complexity. The following factors are accounted for in the proposed model (cf. Hufeisen and Marx (2007: 314):

- Neurophysiological factors, such as general language acquisition capability and age;
- Learner external factors, e.g. the learning environment, type and amount of input;
- Cognitive factors, e.g. language awareness and learning strategies;
- Affective factors, e.g. motivation, anxiety or self-assessment of language proficiency, perceived closeness/distance between languages, attitudes towards languages, target cultures, individual life experiences;
- Foreign language specific factors such as individual foreign language learning experiences, strategies, previous language interlanguages, interlanguage of target language;
- Linguistic factors, i.e. L1, L2, Lx.

According to Hufeisen and Marx, different factors are at play depending on the stage of language acquisition. The first stage, i.e. L1 acquisition, involves only neurophysiological and learner external factors. At the second stage, i.e. L2 acquisition, apart from the factors specific for L1 acquisition, other factors include cognitive and affective variables as well as the first language element from the linguistic factors. The third stage, i.e. L3 acquisition, requires, in turn, foreign language specific factors as well as more linguistic factors (L1 and L2). The next stage of additional language (Ln) acquisition further extends the number of linguistic factors (L1, L2, and L3). The authors claim that the greatest qualitative change occurs between stages two and three, when an array of foreign language specific factors is added into the process of third language acquisition. Due to these qualitative differences in the process of acquisition, TLA should not be subsumed under SLA and should be treated as a separate domain in its own rights.

2.3.2. Hammarberg's Role-Function Model

In his several publications Hammarberg (Hammarberg 2001, Hammarberg 2009, Williams and Hammarberg 1998) attempted to account for the patterns found in third language speech production. Based on the data from a longitudinal case study of a multilingual subject, Sarah Williams, Hammarberg proposed that the languages of a multilingual speaker tend to perform different roles and functions, acting either as instrument or supplier languages. According to Hammarberg (2009: 39), the supplier language functions as supplying material for word formation in the third language and this role was performed in this longitudinal study by the second language (specifically in this case: L2 German for L3 Swedish). On the other hand, L1 English was used in this role rather infrequently, and instead it functioned more as a tool for facilitating communication, e.g. in the form of metalinguistic comments or requests for assistance. This role was referred to as instrumental, and interestingly enough, it was not performed by the second language. On the basis of these observations, Hammarberg concluded that L1 and L2 play different roles in third language acquisition.

An attempt was made to interpret these different roles as reflecting the various activation patterns of background languages. A detailed data analysis demonstrated that an instrumental role was manifested in non-adapted META switches "where the speaker makes use of the background languages in various functions which complement and support the speech production in L3, whereas a supplier role for the background language arises during the L3 production itself in the form of WIPP switches and transfer in word construction" (Hammarberg 2009: 121-122). The status of L1 English as the most readily activated instrumental language for metacomments was further accounted for in terms of personal identification (i.e. maintaining first language identity) and its status as a contact language. The findings indicate a default level of activation for German in the case of the multilingual informant, which resulted in its being regularly active during speech production in L3 Swedish. Therefore, L2 German assumed a default external supplier role, which as Hammarberg argued, was a combined effect of high proficiency, recency of use, typological proximity between German and English, and the L2 status.

2.3.3. De Angelis' Combined CLI

De Angelis' combined cross-linguistic influence (2007: 21) cannot be regarded as a full-fledged model, but rather as a proposal of a conceptualization of an idea. The concept was put forward when discussing different types of transfer, namely, a one-to-one type vs. many-to-one transfer. The former relates to second language acquisition and entails transfer between the source and the target language. The latter pertains to a more complex context of acquisition, which allows for the simultaneous influence of more than one language on the target language. This type of cross-linguistic influence occurs "when two or more languages interact with one another and concur in influencing the target language, or when one language influences another, and the already influenced language in turn influences another language in the process of being acquired" (De Angelis 2007: 21). In the absence of a broadly accepted term for this phenomenon, the author proposed 'combined CLI' as the term of reference. This proposal has been widely tested in third language acquisition research, which is reflected in the literature overview (cf. section 3.2.5).

2.4. Third language acquisition models

Several attempts have been made also to provide theory-based explanatory models designed specifically for multiple language acquisition. The models proposed so far operate mostly in the generative paradigm and were derived from morpho-syntactic empirical investigations. A growing body of empirical evidence is emerging to support the competing models. The following sections will discuss the current three models of multilingual acquisition, including the Cumulative-Enhancement Model (CEM) by Flynn et al. (2004); L2 Status Factor Model by Bardel and Falk (2007, 2012); Typological Primacy Model (TPM) by Rothman (2010, 2011, 2013, 2015).

In an attempt to make logical predictions about sources of multilingual transfer, Rothman (2015: 182) acknowledges 4 possible scenarios: (1) no transfer, (2) absolute L1 transfer, (3) absolute L2 transfer, (4) L1 and/or L2 transfer. No formal model corresponding to the first option, i.e. no transfer, has ever been proposed. Likewise, option two of an absolute L1 transfer has not resulted in a formulation of a model, although some scholars strongly suggested this, including Lazano (2002) and Na Ranong and Leung (2009). The rationale behind the postulate of absolute L1 transfer rested on the assumption of the L1 acting as a filter that blocks any access to L2 properties and the tenet of L1 syntactic representations as the only possible ones (cf. Rothman et al. 2012: 385). So far, the formal models that have been proposed address the last two options of either a prevailing L2 transfer or the interplay of native and non-native transfer.

2.4.1. Flynn et al.'s Cumulative-Enhancement Model

Reflecting on the nature of third language acquisition research, Flynn et al. (2004) suggest that it has the potential to provide novel insights into the process of language learning that surpass those stemming from investigations into first or second language learning. The authors argue that studying L1/L2 acquisition alone "is not sufficient in terms of our understanding of the human capacity for language" (Flynn et al. 2004: 4).

The point of departure for Flynn el al.'s (2004) study was whether the first language (L1) of a multilingual speaker maintains its special privileged role in the acquisition of a subsequent foreign language (L3), as in the case of second language acquisition. The authors aimed to explore if other languages known to a multilingual (i.e. L2 – the first foreign language) can also exert an impact on subsequent language acquisition, thus pointing to a cumulative nature of language learning. The study explored the acquisition of three types of restrictive relative clauses in English as an L3 by L1 Kazakh children and adult subjects with L2 Russian. It focused specifically on the directionality connected with the construction of the Complementiser Phrase (CP) in language specific grammars. The results failed to demonstrate a privileged role of the L1 in the third language acquisition of the selected syntactic structures. It was shown that the patterns of acquisition of the grammatical properties in question in L3 English/L2 Russian/L1 Kazakh matched those reported for L2 English/L1 Spanish rather than those in L2 English/L1 Japanese as predicted by the authors. It suggests that any prior CP development can exert an influence on the development of the CP structure in an additionally learnt language.

The findings are interpreted as providing support for the Cumulative-Enhancement Model for Language Acquisition (CEM) as proposed by Flynn et al. (2004). The major tenet of this proposal is that language acquisition is cumulative in nature, i.e. all previously known languages can have a potential impact on the subsequent language acquisition, thus questioning the validity of a privileged role of the learners' first language (L1) in this process. The model departs from models that view additional language learning as a deficit due to e.g. negative transfer, as it claims that other languages known to a multilingual learner can enhance the development of a subsequent language system whenever it is appropriate, i.e. facilitative. The proposed model excludes redundancy in linguistic representation and maintains that "the accumulated linguistic knowledge necessarily enhances subsequent language learning" (Berkes and Flynn 2012: 144). This means that language acquisition is a collective process that is inherently non-redundant, because the mind tends to avoid repetition. Therefore, the model predicts that any instance of non-facilitative transfer from previous languages would be neutralized or blocked.

Moreover, the proposal contrasts with statistical models for language learning as it assumes that patterns of L3 acquisition are conditioned by already existing representations of linguistic knowledge in the learner's mind. The study poses some questions that remain unanswered, e.g. whether it is actually the latest acquired language (i.e. immediately prior learnt language) that determines the acquisition of the following subsequent language system (Flynn et al. 2004: 13). Further, it remains uncertain if the model is equally applicable to adult as well as child acquirers and whether these acquisition patterns hold both for sequential and (near)simultaneous foreign language acquisition. The interpretation of the results is, however, confounded by the fact that it is the L2 that is the source of transfer in Flynn et al.'s (2004) study, thus one cannot preclude the L2 status effect rather than a facilitative influence of any previously acquired language system. The authors admit that they were confounded by the role of the immediately prior learnt foreign language and left open the question whether it can be a determining factor for a subsequently learnt language.

Berkes and Flynn (2012) provide further evidence in support of CEM by analyzing the CP structure development in different language combinations. The study involved an elicited imitation task aimed at comparing the production of relative clauses by 2 groups of learners (L1 German/L2 English and L1Hungarian/L2 German/L3 English) at different proficiency levels. The results demonstrated an increased facilitation in the subsequent language acquisition with each new language being acquired, reflected in the considerably lower number of S/O conversion errors made by the L1Hungarian/L2 German/L3 English group in comparison to the L1 German/L2 English group. The authors concluded that the last learned language does not exert a negative impact on the subsequent language development and the findings corroborated the assumptions of CEM as an explanatory model for L3 acquisition, specifically at the underlying grammatical structure level that is relevant to CP.

A growing body of literature appears to provide support mostly for a ‘weak’ version of CEM in the sense that transfer in L3 acquisition is not restricted to either a default L1 or L2. Furthermore, several scholars question the motivation for excluding the possibility of non-facilitative transfer. For instance, Rothman (2015: 183) maintains that “having to avoid non-facilitative transfer *a priori* would place an unrealistic burden on limited cognitive resources during the course of forming the emerging L3/Ln system”.

2.4.2. Bardel and Falk’s L2 Status Factor Model

Another theoretical model accounting for sources of cross-linguistic influence in third language acquisition is the L2 Status Factor Model put forward by Bardel and Falk (2007, 2012). The model assumes a privileged position of the second language. In their proposal, the authors refer to earlier claims related to the L2 status factor. The first mention of this phenomenon can be attributed to Meisel (1983) who labelled it as the ‘foreign language effect’, indicating that a previously learnt second language may interfere with the learning of a subsequent foreign language. Based on observations of syntactic and lexical transfer strategies from the second to the third language, he claimed that the conditions for applying this type of transfer do not seem to be identical with those operating in L1-to-L3 transfer. Trying to account for this observation, Meisel proposed an idea that there seems to be “a difference in the neuropsychological basis for storing and processing first and second languages” (1983: 18).

The idea was further elaborated on by Hammarberg (2001, 1993) who defined the L2 status as "a desire to suppress L1 as being 'non-foreign' and to rely rather on an orientation towards a prior L2 as a strategy to approach the L3" (Hammarberg 2001: 36-37). Other scholars have also considered it to be one of the factors determining the sources of transfer, particularly in the acquisition of L3 vocabulary (cf. Cenoz 2001, De Angelis 2005, 2007).

According to Williams and Hammarberg's (1998, 2009) longitudinal case study, L2 German assumed the role of an 'external supplier language' and was used subconsciously, especially in word constructions and insertions in L3 Swedish when the participant was mostly in a foreign language mode. The authors noted that the L1 English, on the other hand, was resorted to consciously, e.g. for metacomments about the L3 oral performance. Hammarberg and Hammarberg's (1993, 2005) study laid the foundations for the L2 Status Factor model.

Bardel and Falk (2007) extended the model to L3 syntax, demonstrating that the L2 is a prevailing source of transfer at the initial stages of the acquisition of a third language, independently of genetic relatedness or any relative typological similarity between the languages involved. The study investigated negation placement in Dutch and Swedish learnt as third languages at the initial stages. While the L2 German group showed no problems with post-verbal negations in the L3, the L2 English group displayed the prevalence for incorrect preverbal negations. The between the groups differences were attributed to syntactic transfer from the respective L2s; however, the pool of subjects was very limited.

The hypothesis that the L2 can supersede the L1 as the main source of transfer in L3 acquisition was based on the assumption of a greater cognitive similarity between the L3 and L2, rather than between the L3 and L1. Consequently, the authors claim that the second language may even block the occurrence of transfer from the L1, even if there is a close typological relation between L1 and L3 and even if such transfer would result in target-like productions in the L3 (Bardel and Falk 2007). Due to similarities between the second language and subsequent foreign languages with respect to such factors as the setting of acquisition, the age of onset, metalinguistic knowledge and learning strategies, the L3 learners tend to classify differently their native vs. foreign languages, and thus, co-activate non-native languages in subsequent language acquisition.

The L2 status factor model was further tested with L3 learners at an intermediate level and the findings appeared to confirm the initial hypothesis (Falk and Bardel 2011). This study explored the placement of object pronouns in L3 German based on a grammaticality judgement task in two mirror design groups with L1 French/L2 English and L1 English/L2 French. There was a significant difference in the response patterns of the two groups, and the accuracy rates for the recognition of ungrammatical sentences in German reflecting either English or French word order clearly indicated a negative transfer effect from the respective L2s. When accounting for their findings, the authors pointed to several differences between the native and non-native languages in terms of the manner and route of acquisition, resulting in the cognitive dissimilarity between L1 and L2/Ln, thus explaining the privileged role of the L2 as a transfer source in learning the L3.

In another paper, Bardel and Falk (2012) resorted to Paradis' psycholinguistic model of bilingualism (2004), thus providing a neurolinguistic basis for the predictions of the L2 Status Factor model. The authors tried to account for the L2 status impact by pointing to the distinction between the declarative and the procedural memory (cf. Paradis 2004, 2008), which may explain greater cognitive similarity between L2 and L3 rather than between L1 and L3. Adopting a neurolinguistic approach to L3 learning, the authors ascribed to Paradis' claim (2008: 344) that "all late-learned languages (L2, L3, Ln) are sustained to a large extent by declarative memory. As such they are more likely to manifest dynamic interference from one another than from the native language(s)". According to Paradis (2009), implicit linguistic competence and explicit metalinguistic knowledge that are both responsible for our verbal communication, remain neurolinguistically distinct. Moreover, these two types of knowledge have various memory sources and cerebral representations, with the implicit linguistic competence being sustained by the procedural memory involving non-conscious representations that rely on the right cerebellum; whereas the explicit metalinguistic knowledge is sustained by the declarative memory that has conscious representations relying mostly on the hippocampal system (Paradis 2009: 139). Paradis further explains the varying routes of the acquisition of different language components in the native language vs. foreign languages by juxtaposing the reliance on procedural memory in the implicit acquisition of L1 phonology, morphology, syntax and lexicon as opposed to explicit learning of these components, sustained by the declarative memory in the L2 and any subsequent foreign language. With the development of L2 proficiency there is a gradual shift from the reliance on explicit metalinguistic knowledge towards more implicit competence and, consequently, some automatization of the foreign language performance. Nonetheless, as maintained by Paradis (2009: 101), "[s]ome aspects (e.g. morphology or syntax) may be automatized while others (e.g. phonology) continue to be controlled – one more illustration of the neurofunctional modularity of the components of linguistic competence". He further claims that the two types of knowledge may complement each other working in parallel, yet they do not interact with each other.

An interesting distinction is made between the status of function vs. content words in L1. While the former are acquired implicitly, stored in the procedural memory and processed as part of syntax, the latter are ascribed the status of vocabulary, which is sustained by declarative knowledge. This may explain different transfer patterns in L3 for these two categories of words; with L1 and L2 both being transfer sources for L3 content words and L2 transfer prevailing in the case of function words (cf. Bardel and Falk 2012).

Bardel and Falk (2012) conclude that the L2 status factor is a strong predictor of both negative and positive transfer from L2 to L3, based on the neurolinguistic and cognitive similarity between these languages as opposed to the first language. Non-facilitative transfer is predictable, in line with Rothman's TPM, although the rationale behind this claim is different in both models.

Recapitulating, the L2 Status Factor Model appears to be a strong hypothesis that offers straightforward predictions which can be tested regardless of language repertoires (L1, L2, L3). However, it cannot be ruled out that L2 transfer could be caused by structural or some other factors. Therefore, Rothman (2015: 182) in his critical evaluation of this model questions, to some extent, the validity of the presented explanation based on Paradis' theory by asking how factors of structural similarity could bypass the filter imposed by the differences between L1 and L2 in the mental storage and the type of representation. Still, the L2 Status Factor model invites further testing to verify its claims.

2.4.3. Rothman's Typological Primacy Model

Another factor that has attracted a lot of attention in the L3 literature is the typology factor. However, there seem to be different understandings of the term, as some scholars view it as overall language relatedness, whereas others as the similarity of particular structures between the source and the target language, irrespective of typological relatedness (cf. De Angelis 2007). Further, it can be postulated on purely linguistic grounds or approached from the learner's perspective as psychotypology (Kellerman 1983). Many studies demonstrate that the most similar and/or the most closely related background languages are selected as sources of transfer in L3 acquisition (e.g. Cenoz 2001, Ringbom 1987).

The basic tenets of the Typological Primacy Model (TPM) were put forward in Rothman (2010) and further modified and expanded in Rothman (2011, 2013, 2015). The TPM is an acquisition model constructed within the generative framework and it relies on formal linguistic theory to specify its claims.

In his 2010 study, Rothman investigated patterns of acquisition of L3 Brazilian Portuguese by two groups of learners with mirror design language pairings: L1 English/L2 Spanish vs. L1 Spanish/L2 English. The study tested word order restrictions in transitive vs. intransitive verbs in different types of sentences. The results demonstrated unambiguously that Spanish was the source of transfer for L3 Brazilian Portuguese, irrespective of the status of Spanish either as the L1 or the L2, thus pointing to the primacy of the typological similarity between the languages involved. The findings were interpreted as counterevidence against the predictions of the CEM and L2 status model. Thus, the relative structural similarity between the L3 and one of the previously acquired languages proved to be a determinant of the L3 transfer.

What the TPM has in common with the CEM model is the assumptions of multiple sources of transfer as well as access to both the L1 and L2 at the L3 initial state. One of the major differences between the models is, however, that the TPM predicts also the possibility of non-facilitative transfer. Furthermore, although both models consider multilingualism to be conditioned by the cumulative influence of previously acquired language systems, they provide varying accounts for the sources of transfer into the L3. The CEM envisages the transfer to be facilitative or neutral in nature, whereas according to the TPM, the source language for transfer is determined on the basis of structural similarity at the underlying level of linguistic competence between L3 and L1 or L3 and L2, its driving force being cognitive economy. Furthermore, The TPM does not negate the validity of the L2 Status Factor Model, yet it postulates the existence of an internal parser with whose assistance the learner evaluates the typological similarity of previously known languages to the subsequent language being acquired and, consequently, selects a language system that is more typologically similar as a source of transfer to the L3.

The basic tenets of the Typological Primacy Model were put forward in Rothman (2011) as modifications of Flynn et al.'s (2004) CEM. The model stipulates that the "Initial State transfer for multilingualism occurs selectively, depending on the comparative perceived typology of the language pairings involved, or psychotypological proximity" (Rothman 2011: 112). Empirical support comes from the study on noun raising in two groups: L1 Italian/L2 English/L3 Spanish and L1 English/L2 Spanish/L3 Brazilian Portuguese. The data demonstrated that L3 learners transferred the necessary syntactic knowledge either from L1 Italian or L2 Spanish and that there were no significant differences in the performance of both groups. The author claimed that the selection of the source of transfer was based on the overall typological proximity between the language systems involved.

In yet another article, Rothman (2013) tries to justify the claims related to restrictions that the multilingual mind places on the selection of transfer. First of all, he maintains that the assessment of typological proximity needs to take place very early in the L3 acquisition process. Further, he claims that once the selection of L1 or L2 has been made, full transfer of this language system occurs, as the potentially feasible property-by-property transfer would be gradual, and thus slower and less economical. Based on the knowledge that the executive control system operates differently in the bilingual mind compared to that of a monolingual, and that inhibition processes are necessary to suppress the activation of the other system, it seems obvious that the cognitive burden is considerably more complex in multilingual acquisition. Therefore, Rothman believes that the proposed claims of early typological assessment and complete transfer constitute the most efficient strategies in L3/Ln acquisition. Rothman (2013) provides a hierarchy of linguistic cues from L3 input, on the basis of which the parser determines the structural proximity, which represents the following order of impact: lexicon > phonological/phonotactic cues > functional morphology > syntactic structure.

In his newest article (2015), Rothman elucidates the assumptions of the Typological Primacy Model by providing linguistic and cognitive motivation for the model and considering such factors as the timing of acquisition and proficiency. The major tenet of the TPM is that transfer in the L3 is determined by structural proximity between the L3 and the L1 and/or the L2. In his latest proposal the author stipulates how typological (structural) proximity is determined unconsciously on the basis of linguistic cues from the L3 input. L3 input is used by the parser at the initial stages of acquisition to determine the source of holistic transfer from one of the previously available language systems (i.e. L1 or L2). Rothman clarifies his understanding of typological or structural similarity (both terms used interchangeably) by referring to linguistic properties which overlap cross-linguistically at the level of mental representation either at the lexical or grammatical level (Rothman 2015: 179). Interestingly, in his view, typological similarity can be either actual or perceived. He further claims that such structural similarity "is assessed and determined subconsciously by the linguistic parser very early in the L3 process based on an implicationally hierarchical continuum of linguistic cues" (Rothman 2015: 179). This leads to a selection of either the L1 or the L2 as typologically closer to the third language as the only source of complete transfer to L3, following a brief transitory stage when both of the systems are available.

Rothman's claims are motivated by the principles of general cognitive economy, according to which the human mind is predisposed towards the least effort when engaging in a cognitive task. According to the author of the TPM, this inclination is manifested also in the transfer patterns in L3 which are driven, among others, by processing cost reduction. In TPM, L3 transfer is argued to work holistically (i.e. full transfer) rather than on a structure-by-structure basis, and consequently it may be both facilitative or non-facilitative in nature as evidenced e.g. in Rothman and Cabrelli Amaro (2010). In this investigation into patterns of acquisition of pronominal subjects, the performance of two L3 groups was compared, i.e. L3 French vs. L3 Italian, which shared their background languages, i.e. L1 English and L2 Spanish. According to the TPM predictions, L2 Spanish would be selected as the source of transfer in both groups due to the typological proximity and the study confirmed these predictions. In the case of L3 Italian the transfer from Spanish was found to be facilitative, as both languages have null subjects; however, for L3 French the transfer was non-facilitative since French (and English) are not null-subject languages.

Alluding to other theoretical claims on the foundations of language (e.g. Jackendoff 2002), Rothman sees the manifestation of cognitive and linguistic economy in the dynamic reality of multilingual acquisition (Rothman 2015: 180). Similarly to Flynn et al. (2004), Rothman points to special insights offered by research on third language acquisition and claims that studying multilingual transfer patterns "permits a unique window into language and cognition in ways that cannot be seen in monolingualism or bilingualism" (2015: 181). He maintains that investigations into the acquisition of a third language contribute towards a more precise understanding of how language is represented in the mind. In his most recent work, Rothman (2015) tries to address previously unanswered questions connected to the cognitive and linguistic factors that determine structural similarity as such an influential factor in L3 transfer. He further intends to test if TPM can predict differences in the nature of transfer in simultaneous vs. sequential bilingualism or transfer from a not fully developed second language. Rothman offers a general definition of transfer, not related to any specific research paradigm, according to which "transfer refers to performance behaviour in a target language that can be reasonably linked to influence from previous linguistic experience" (2015: 181). He further clarifies that from the point of view of the generative approach to acquisition, he differentiates transfer at the level of underlying competence from other sources of learned linguistic and metalinguistic knowledge.

All in all, Rothman has attempted to tease apart the factors of typology and L2 status in a number of studies. The general conclusion is that typology can overrule the L2 status factor when the L1 is closely related to the L3. Rothman has demonstrated further empirical evidence in favour of the TPM stemming from a growing body of research in the field. Several recent studies on various language repertoires were found to provide some support for the TPM, e.g. Kulundary and Gabriele 2012, Wrembel 2012.

The limitation of Rothman's studies conducted so far is that they focus on the language triad including English, Spanish and Brazilian-Portuguese, the last two being closely related languages. Moreover, Slabakova (2015) questioned some of the assumptions of the TPM in her recently proposed Scalpel Model. Although she also maintains that transfer can be facilitative as well as detrimental and that it can occur from either L1 or L2, yet the claims are based on grammatical considerations rather than typological proximity. Slabakova opposes the idea of there being a wholesale transfer which blocks some linguistic information as less economical from the point of view of language neural organization, suggesting instead a property-by-property transfer. She further claims that there are other factors such as processing complexity, misleading input, L2 proficiency or construction frequency that condition the process of L3 acquisition; however, these suggestions remain open to further empirical testing.

All in all, the current theoretical models of third language acquisition offer quite conflicting explanatory accounts of the phenomenon in question. While there seems to be agreement as to the interplay of native and non-native sources of cross-linguistic influence, the predicting factors underlying its strength and directionality remain still inconclusive. Recapitulating, I would like to reiterate Rothman's (2015: 189) conclusion that in the contemporary world in which multilingualism prevails over monolingualism as "the default state of linguistic knowledge", we do need to be able to account for multilingual acquisition, and therefore the existing models have to be further verified and developed.

2.5. Theoretical conceptualization of present studies

Scholars involved in research on multilingualism have aimed to develop unified theories of third language acquisition; however, the competing current models are not yet able to account fully for the complexity of the phenomenon. The authors of particular models univocally call for further empirical testing of the proposed hypotheses to be carried out in various subdomains. The aim of the present contribution is therefore to inform the ongoing debates on the specificity of phonological acquisition in the third language by conducting a series of parallel studies into the field.

No specific model for the acquisition of third language phonology has been proposed so far. A question arises whether the existing models of acquisition of second language speech or the current models of third language acquisition are powerful enough to explain the processes involved in the phonological acquisition of a third language.

The present contribution is intended to test the assumptions put forward in the models of third language acquisition discussed in Chapter 3, particularly the ones related to the sources and directionality of cross-linguistic influence which is the focal research concern in the present empirical investigation. The series of studies was designed in such a manner as to substantiate a potential mutual impact of the languages from multilinguals' repertoire, and particularly, to develop explanations for the role of native vs. non-native languages in shaping phonological acquisition patterns in the third language.

2.5.1. Research hypotheses

With the view to the aims specified above, the following research hypotheses were developed:

H1: Cross-linguistic influence is conditioned by the L1 effect
H2: Cross-linguistic influence is conditioned by the L2 status
H3: Cross-linguistic influence is conditioned by the typological proximity
H4: Cross-linguistic influence is conditioned by an interplay of the above factors.

The proposed hypotheses stem from the assumptions of the formal models of third language acquisition, including the Cumulative-Enhancement Model (Flynn et al. 2004); L2 Status Factor Model (Bardel and Falk 2007, 2012) and Typological Primacy Model (Rothman 2010, 2011, 2013, 2015). They aim to test which of these general models is the most applicable to account specifically for the process of acquisition of third language phonology.

In order to verify the main research hypotheses, several specific research questions were posed as specified in the conducted series of studies:

1) Study I on the perception of foreign accentedness – see 4.2.1
2) Study II on the acoustic measures of voice onset time – see 5.2
3) Study III on metaphonological awareness – see 6.2.1

Further, specific research questions were formulated with respect to the planned comparative analyses between the study groups involved in the empirical testing, i.e. groups A, B, C, and D (see section 4.1 for a detailed description thereof). The following four groupings were proposed as a basis for a comparative analysis of data generated in the series of conducted studies and the subsequent testing of the main research hypotheses. The groupings were based on a mirror design principle, in which one of the languages was a stable variable (i.e. L1 Polish) and the remaining three languages (i.e. English, French and German) were grouped interchangeably as either the L2 or L3, thus creating varying testing conditions. The group pairings included conditions with the same L2, with the same L3, with typologically related vs. unrelated language pairs, and with the same vs. different second languages, as presented below:

1) A vs. B – a group pairing with the same L1 and L2, but different L3s
   A: L1 Polish, L2 English, L3 French vs. B: L1 Polish, L2 English, L3 German.

2) C vs. D – a group pairing with the same L3 and L1, but different L2s
   C: L1 Polish, L2 German, L3 English vs. D: L1 Polish, L2 French, L3 English.

3) B vs. C and A vs. D – typologically related and unrelated pairs of groups
   B: L1 Polish, L2 English, L3 German vs. C: L1 Polish, L2 German, L3 English,
   A: L1 Polish, L2 English, L3 French vs. D: L1 Polish, L2 French, L3 English.

4) A vs. B and C vs. D – pairs of groups with the same L2 and different L2s
   A: L1 Polish, L2 English, L3 French vs. B: L1 Polish, L2 English, L3 German,
   C: L1 Polish, L2 German, L3 English vs. D: L1 Polish, L2 French, L3 English.

The comparative analysis of the selected L3 phonetic performance measures was designed to verify specific research questions and hypotheses assigned to particular grouping conditions as stipulated below (see also Table 1):

1) Group pair comparison with the same L1 and L2, but different L3s.

Does the performance on three measures vary for different L3 groups irrespective of the fact that the L2 is the same? If so, this may be due to the influence of typology as the L2 status is being kept constant, and there is a different typological proximity relation between the L2 and L3 in the two groups (i.e. closer proximity in Group B, more distant in A). If the performance of the two groups is the same, this may be due to the L2 status or the L1 effect. If not, the difference in the performance may be attributed also to the impact of different features of the respective L3s.

2) Group pair comparison with the same L1 and L3, but different L2s

Is the L3 performance in English comparable irrespective of the fact which L2 a given group has in its linguistic repertoire? If there is a difference in phonetic performance between the two groups, it may be attributed to the L2 status effect that exerts an impact on the L3 phonological performance. The L2 status effect may be further reinforced by the influence of typology, i.e. we may hypothesize a facilitative impact of typological proximity between L2 and L3 in Group C, which should give it an assumed advantage over Group D.

3) Comparison between typologically related and unrelated pairs of groups

If the first two groups which have typologically related L2 and L3 behave similarly and, at the same time, differently from the remaining two groups, which are typologically unrelated, it may point to the impact of the typological closeness or distance as factors determining their performance in the L3, which may overrule the L2 status or L1 effect. If there are no differences in performance reported between these two pairs of groups, it will fail to provide evidence for the conditioning role of typology. Conversely, it may indicate the L1 effect on the L3 phonological performance, since the L1 Polish is the only shared linguistic background.

4) Comparison between pairs of groups with the same L2 and different L2s.

If the first two groups sharing the L2 behave the same and the remaining two groups do not, it may be interpreted as a strong indicator of the L2 status. If there are no differences between the performance between these two pairs of groups, the role of the L2 status as a predicting factor in third language acquisition will not be confirmed as a dominating factor.

Table 1. Testing assumptions for four grouping conditions

Groupings	Conditions	Testing assumptions
A vs. B	same L2, different L3s	L2 status constant verify L3 impact or typological advantage (B)
C vs. D	same L3 different L2s	L3 characteristics constant L2 status effect or typological advantage (C)
B, C vs. A, D	typologically related vs. unrelated pairs	typological distance irrespective of L2 status
A, B vs. C, D	pairs with same L2 vs. different L2s	L2 status or L1 effect

2.5.2. Predicted scenarios

In an attempt to predict the selection of a source language for phonological cross-linguistic influence in L3 acquisition, five potential general outcomes were hypothesised, as stemming from the selected models of foreign language acquisition:

(1) native L1 Polish influence would override the non-native influence resulting in an L1-accented performance in the L3;
(2) non-native influence of L2, or the so-called ‘L2 status’, would be a prevailing source of cross-linguistic influence leading to a perceived L2-accented speech in L3;
(3) both the native and non-native languages would have an impact on the perceived foreign accent in the L3, thus substantiating the assumption of a combined cross-linguistic influence;
(4) other intervening variables such as typological proximity can account for the observed patterns of CLI;
(5) patterns of CLI escape any categorizations from 1-4.

The first scenario follows indirectly from the general assumptions of second language acquisition models of speech, according to which the first language acts as a filter to second language perception and production of speech (e.g. Flege 1995, Best 1995, Dziubalska-Kołaczyk 1990). When treating the acquisition of third language phonology as a mere extension of the second language, one may postulate that the same principles should underlie this process thus the application of SLA models of speech should be justifiable. However, the above mentioned theories do not account for the existence of more than one foreign languages in the repertoire of the learner, thus the application of these models can be only indirect or tentative. On the other hand, the hypothesised L1 dominance in L3 acquisition is also in line with some theoretical claims put forward in the field of third language acquisition, in accordance with which the sensimotor routines established in the first language acquisition are also dominant in any subsequent language acquisition (cf. Ringbom 1987).

The second scenario promoting the second language as the major source of cross-linguistic influence in the L3 acquisition stems from the tenets of the L2 Status Factor Model by Bardel and Falk (2007, 2012) (see section 2.4.2 for a thorough discussion thereof) as well as the main claims of Hammarberg's (2009) Role Function Model (cf. section 2.3.2). The models assume a privileged role of the L2 rather than the L1 in the subsequent language acquisition, based on a greater cognitive similarity between the non-native languages related to the setting and age of acquisition as well as metalinguistic knowledge. The special status of the second language in L3 acquisition is accounted for by a neuropsychological difference in the storing and processing of the native vs. non-native languages (cf. Meisel 1983, Paradis 2008, 2009).

According to the third scenario, both the first and second languages exert an impact on the acquisition process of the third language, which corresponds to some assumptions of the Cumulative Enhancement Model (Flynn et al. 2004) (see section 2.4.1 for a detailed discussion) and is in line with De Angelis' (2007) proposal of a combined cross-linguistic influence (cf. 2.3.3). Both models stipulate that language learning is cumulative in nature and that the accumulated linguistic knowledge enhances subsequent acquisition, while acknowledging, at the same time, that the L1 does not necessarily maintain a privileged position in this process.

In the fourth scenario stipulates that it is other intervening variables such as typological proximity which can determine the actual source of CLI and account for the observed patterns of linguistic interactions. It is based on the tenets of the Typological Primacy Model (Rothman 2010, 2011, 2015), in accordance with which multiple potential sources of transfer are determined by the structural similarity between previously known languages and the one being currently acquired (cf. section 2.4.3). The claims of this model are embedded within a formal theoretical framework of generative linguistics and are further motivated by the principles of general cognitive and linguistic economy.

The final scenario provides a negative rendition of the previously discussed four potential outcomes. It envisages a situation in which the observed patterns of CLI would escape any categorizations hypothesised in points 1 to 4. The verification of the specified research hypotheses and the predicted scenarios will be discussed in Chapter 7, section 7.3.

Chapter 3

Studying L3 phonology; an overview of research

3.1. Introduction to L3 phonology

In the present era of multilingualism "the study of L3 phonology is becoming a field that is both theoretically relevant and practically necessary", as emphasised by Cabrelli Amaro (2012: 54). Several scholars indicate, however, that research on the phonological acquisition of a third language (L3) is still in its infancy and has not been a primary focus of studies to date (cf. Cabrelli Amaro 2012, Gut 2010), although other domains of L3 acquisition such as lexicon or morphosyntax have received a much wider coverage. In an early publication Ringbom (1987: 114) indicated that "the effect of grammar and phonology […] is accorded much less space and importance" and phonology is hardly ever studied from the L3 perspective. This remark was seconded later on by Hammarberg (1997) who also expressed his concern about the lack of research into L3 phonological acquisition. However, in a state-of-the-art overview article Cabrelli Amaro (2012) concludes that although third language phonology has been an understudied domain, it has recently experienced a significant development.

The 21st century has witnessed an upsurge of interest in the area, reflected in a growing number of related publications (e.g. a Special Issue on transfer in L3 phonology published by the International Journal of Multilingualism in 2010, guest edited by Wrembel, Gut and Mehlhorn) and a greater presence of research on L3 phonology at international conferences including New Sounds, EuroSLA or the International Conference on Third Language Acquisition and Multilingualism. Moreover, special workshops were organised devoted to L3 phonology including a satellite workshop of the International Congress of Phonetic Sciences (ICPhS) on Phonetics and Phonology in Third Language Acquisition held in Freiburg, Germany in 2007 or a workshop on Advances in the Investigation of L3 Phonological Acquisition organised at the Societas Linguistica Europaea (SLE) conference in Poznań, Poland, in 2014.

The existing body of research on third language phonological acquisition demonstrates an inherent complexity of this growing field. Further research is still necessary to offer a more comprehensive picture of this understudied domain. To this end, the present contribution aims to bridge the gap in the existing literature by providing more insights into problem areas in third language acquisition of phonology.

3.2. Overview of research on L3 phonology

The present overview of research into the acquisition of third language phonology is partially based on previous shorter surveys presented in Wrembel (2012, 2014, 2015); however, this version is the most exhaustive one. The review of the literature ranges from the earliest accounts of multilingual case studies (e.g. Chamot 1973, Rivers 1979), through a seminal study by Hammarberg and Hammarberg (1993), to the most recent empirical investigations into the acquisition of L3 phonology, the majority of which has been published in the last decade.

The following sections are intended to review the limited body of research on L3 phonology, to discuss their major findings and limitations as well as point to methodological considerations. The overview is organised according to the chronology and scope of research, starting with the earliest accounts of phonological acquisition in multilinguals dating back to the 70s through more controlled experimental studies from the 80s and 90s, including the longitudinal case study by Hammarberg. It then proceeds to discuss investigations into the phenomenon of cross-linguistic influence organised according to the evidenced sources of transfer in L3 phonology, with separate subsections testifying L1 transfer dominance, L2 status effect and combined transfer sources. Further, more recent investigations into L3 phonology are presented, with the scope of research ranging from perceptual discrimination and categorization tests on vowel quality and quantity to production tasks investigating speech rhythm and vowel reduction. Separate sections are devoted to research on selected phonetic/phonological properties including voice onset time (VOT), accentedness ratings and metaphonological awareness, all of which remain focal for the present dissertation. Finally, some existing interdependency studies are outlined and general methodological considerations related to research on L3 phonology are discussed.

3.2.1. Early case studies

The earliest evidence of scientific interest in the area of phonetic transfer in multilingual acquisition comes from the work by Chamot (1973), Rivers (1979) and Singh and Carroll (1979).

Chamot (1973) reported on a case study of a French-Spanish bilingual acquiring English as a third language. The author found evidence for a 'double interference' of two language systems that concurred in influencing the phonetic system of the target language. In other words, both French and Spanish exerted an influence on the learner's phonology in English acquired as the third language.

Similarly, Rivers (1979) provided further support for the phonetic influence going beyond L1 transfer on the basis of a diary study of a multilingual informant with L1 English, L2 French, L3 Spanish, and L4 Italian and German, with the languages being classified as L1/L2/L3/L4 on the basis of their chronological order of acquisition. The study showed that the Romance languages (i.e., French and Italian), acquired as foreign languages, influenced the learner's L3 Spanish more strongly than her native English. Interestingly, Italian was a previously learnt language that was nearly forgotten, yet it became reactivated in subsequent language learning. This led Rivers to suggest that language typological distance and phonetic resemblance constitute factors which facilitate cross-linguistic influence from non-native languages.

Further evidence of non-native phonological transfer was offered by Singh and Carroll (1979) who investigated native speakers of non-European languages with L2 English who learnt French as their third language. The results showed a significant influence of L2 English on the subjects' L3 French speech.

The major limitation of these three earliest accounts of the acquisition of third language phonology is that they were mostly impressionistic studies and did not rely on experimental designs. Nonetheless, the qualitative observations stemming from these multilingual case studies inspired further interest in the area and eventually led to more controlled investigations into cross-linguistic influence in L3 phonology.

3.2.2. Hammarberg's study

One of the most frequently quoted studies on third language phonological acquisition is a pioneering longitudinal case study by Hammarberg that seems to have laid the foundations for current investigations into L3 phonology. The case study examined language development in a multilingual informant, Sarah Williams, whose language repertoire included L1 English, L2 German and L3 Swedish as well as other less proficient foreign languages, and various aspects of this process were reflected in a number of publications (e.g. Hammarberg 2001, 2009; Hammarberg and Hammarberg 1993, 2005; Williams and Hammarberg 1998).

The phonology-oriented part of the investigation involved accent judgements conducted on samples of L3 Swedish performed by 3 native listeners at different stages of interlanguage development as well as an auditory and acoustic analysis conducted by Hammarberg and Hammarberg (1993, 2005). The results pointed to a stronger interference of the non-native language, that is, L2 German rather than L1 English at an early stage of the L3 acquisition when the informant's performance in L3 Swedish was found to be strongly L2-accented. Nonetheless, L2 interference was shown to diminish in the course of time and the native English accent dominated at a more advanced stage of L3 learning, thus leading to a more L1-accented oral production in Swedish. The findings from the perceptual accent judgement task were corroborated by an auditory analysis of the informant's L3 Swedish speech performed by the authors.

A particular novelty of Hammarberg's approach consisted in his investigation of the resetting of an articulatory basis in the acquisition of a new language. It was based on earlier interest in articulatory settings or voice quality (e.g. Honikman 1964, Laver 1980, 1994), i.e. a general articulatory posture in neutral as well as active positions of the speech organs that are characteristic for particular languages. Both the perceptual ratings as well as the auditory analysis of the multilingual informant's performance in L3 Swedish at different stages of interlanguage development demonstrated more German-like general colouring of her speech at the initial stages of acquisition as opposed to a more English-oriented articulatory setting at a more advanced stage of L3 learning. To further test the differences in phonatory settings, Hammarberg performed acoustic measurements of fundamental frequency distribution. However, a comparative analysis of F0 frequency and F0 range between the L2 German and L1 English extracts read by the subject proved to be inconclusive. Sarah Williams' ability to switch between English and German settings was interpreted as an indication of the articulatory re-setting process that occurred during second language learning.

The study also showed a task-related variability since imitation and repetition tasks involved in the research design generated a less L2-accented performance in L3 Swedish, whereas L3 reading and free production, i.e. tasks which were subject to less conscious monitoring, resulted in a stronger L2 interference in the L3 output. According to Hammarberg (2009), the rationale behind the prevalence of the L2 accent at the initial stages of L3 acquisition could be attributed to the high proficiency in L2 German and the recency of its use, on the one hand, as well as a conscious avoidance strategy on the part of the subject not to sound as a native English speaker, on the other.

Hammarberg concluded that although L1 influence is a basic constraint on articulation due to the established neuro-motor routines, "the reliance on L2 settings is a coping strategy which the learner resorts to at an initial stage when the phonetic form of L3 is too unfamiliar to master, and abandons when proficiency in L3 increases. This coping strategy is seen to override the basic constraint temporarily" (2009: 84). This study laid some foundations for the acknowledgement of varying sources of cross-linguistic influence in L3 phonological acquisition including both native and non-native languages.

3.2.3. L1 transfer prevalence

A relatively limited number of studies have investigated the phenomenon of cross-linguistic influence in the acquisition of third language phonology. Initial experiments conducted on L3 phonology in the 1980s pointed to the native language as the primary source of negative transfer (Llisterri and Poch 1987, Pyun 2005, Ringbom 1987), following the traditional approach to foreign language acquisition, widely attested in SLA research, that saw transfer as a one-to-one phenomenon from the L1 to the target language.

Along these lines, Ringbom (1987) maintained that L1-accented speech is evident even in advanced L3 learners, particularly in the area of intonation, and that instances of transfer from the second language are rather infrequent. A similar conclusion was reached by Garcia Lecumberri and Gallardo del Puerto (2003), whose study on Spanish/Basque bilinguals learning L3 English also demonstrated that there was an L1 influence.

Llisterri and Poch-Olivé (1987) provided further support for this claim in their studies of Spanish/Catalan bilinguals who acquired French or English as their third languages. In the investigation of L3 French vowels, the frequency of the first two formants was compared to the performance of monolingual Catalan speakers. With respect to the measured durations of L3 French vowels, the results indicated that both monolingual and bilingual speakers did not make significant differences between long and short vowels. As far as fricative consonants in L3 French are concerned, no significant differences were found either between the two groups with respect to the measured frequency and intensity of acoustic energy or the duration of the consonants. The authors concluded that they found no evidence of L2 interference on the oral production in L3 and the existing transfer could be satisfactorily explained by means of the acoustic features of the subjects' L1 sounds. It appeared that bilingual speakers learning a third language behaved in the same way as their monolingual Catalan counterparts as they tended to follow their first language distributional patterns in the acoustic vowel space. The study, however, suffered from some limitations including a small sample of subjects and the selection of language pairings, i.e. the fact that Catalan and Castilian are closely related languages, and that Catalan is closer to French than Castilian Spanish, which may have influenced the results.

The results of the present author's previous exploratory investigations into foreign accentedness in a third language also indicated some evidence of L1 transfer (Wrembel 2012a). However, when compared to other related studies (e.g. Wrembel 2012b), the findings of L3 accentedness ratings could be interpreted as not just pointing to L1 transfer or L2 status as predictors of the source of cross-linguistic influence but rather as the interplay of several factors, including typological proximity between the phonological systems involved.

In the early studies, the L1 transfer has been widely attested as the major factor affecting third language acquisition, while the non-native sources of CLI have not been recognised as significant. Notwithstanding this, more recent research has focused on the role of the first acquired foreign language (L2) in subsequent phonological acquisition, and the findings attesting the L2 status effect will be discussed in the following section.

3.2.4. L2 status effect

A number of more recent studies have shown that cross-linguistic influence in L3 phonological acquisition is not limited only to the native language and that non-native languages tend to exert a considerable influence on the developing L3 system (e.g. Gut 2010; Llama et al. 2010; Tremblay 2007; Wrembel 2010). Some scholars (e.g. Bardel and Falk, 2007) working in the TLA framework even claim that the influence of the second language (L2) prevails as the source of transfer and overrides L1 transfer thus testifying to the existence of the so called 'foreign language effect' or 'L2 status' in L3 phonological acquisition.

Hammarberg and Hammarberg's (1993, 2005) study described in one of the previous subsections (see 3.2.2) clearly evidenced a leading role of the non-native language in shaping L3 phonology, especially at the onset of acquisition. The authors tried to explain this phenomenon in terms of the reactivation of L2 learning mechanisms triggered by the cognitive similarity involved in the process of non-native language learning as opposed to naturalistic first language acquisition. A detailed account of the varying roles of L1 and L2 in subsequent language acquisition is offered in Hammarberg's (2001) Role Function Model.

Trembley (2007) reported facilitative transfer from L2 French in the acquisition of VOT patterns by L1 English learners of L3 Japanese (see section 3.2.7 for further details). Similarly, Kamiyama (2007) observed an L2 influence, although of a non-facilitative nature, in his study on the acquisition of L3 French vowels by L1 Japanese learners with L2 English. Moreover, Llama et al. (2010) demonstrated that cross-linguistic influence from the second language was a stronger predictor of VOT patterns in L3 Spanish in two groups of French-English bilingual participants (see section 3.2.7 for a detailed discussion).

Also Wrembel (2010) aimed to investigate the degree of influence of the native and non-native languages on L3 phonological acquisition by means of perceptual judgements of a foreign accent in L3. This study demonstrated that the trilingual speakers of L1 Polish, L2 German and L3

English showed evidence of L2-accented speech in their L3 performance at the early stages of acquisition, yet this tendency diminished with the advancement of language proficiency. The results were consistent with Hammarberg and Hammarberg's claims (1993, 2005); nonetheless, they should be interpreted as a cumulative effect of several variables, including the foreign language effect or L2 status, typological proximity (with English and German being Germanic languages, one may expect more cross-linguistic influence from German than from Polish) as well as the recency and intensity of L2 German use.

Several other studies that found evidence of L2 influence on phonetic performance in L3 pointed to typological closeness or perceived proximity (i.e. psychotypology) between the language systems involved rather than the L2 status alone as the driving force behind this process. Further examples of related research are provided in section 3.2.6 on the most recent investigations into the area.

3.2.5. Combined transfer

The literature on third language phonological acquisition has also reported multiple sources of transfer, thus testifying to a combined transfer phenomenon, as proposed by De Angelis (2007). Sources of combined transfer are understood to include any previously acquired language systems as well as language universals, i.e. phonological processes and rules occurring systematically across natural languages (cf. Cabrelli Amaro 2012).

A study by Benrabah (1991) provided support for combined sources of phonological transfer in Algerian Arabic/French bilinguals acquiring English as their L3. Noteworthy, the source of transfer was found to be determined by the complexity of the existing subsystems, i.e. vowels were transferred to L3 English from L2 French, which features a more complex vocalic repertoire, whereas consonants were transferred from L1 Arabic due to a larger complexity of the consonantal set.

Pyun (2005), on the other hand, maintained that he found evidence of phonological knowledge in speakers of L3 Swedish from four different sources including phonological rules and categories from L1 Korean, L2 English, L3 Swedish as well as interrules between these phonological systems. The examples of processes from more than one source provided by Pyun clearly point to a combined transfer effect. The author accounted for different phonetic realisations of L3 lexical items by means of this interplay of various rules, but it remains unclear if more systematic generalisations can be drawn from this.

Other studies that focused on selected features also provided further support for combined transfer. For instance, Wunder (2010) in her investigation of VOT patterns in L3 Spanish found both the influence of L1 German and L2 English. She reported the transfer of hybrid VOT values from L2 to L3 as well as an underlying L1 effect. Similarly, Blank and Zimmer (2009) found evidence of hybrid vowel productions in L3 English by speakers of L1 Brazilian Portuguese and L2 French. The performed acoustic analysis of fundamental frequencies F1 and F2 as well as durational measurements, demonstrated that L3 English vowels were influenced both by L1 and L2 phonological systems. However interesting these findings are, a visible weakness is that the L1 was not tested in either of the studies and that monolingual reference values were used instead for comparison against acoustic measurements in L2 and L3. The lack of L1 data collected from multilingual subjects could be criticised based on evidence from SLA studies, according to which the L1 phonological system undergoes modifications under the influence of the first acquired foreign language (e.g. Flege 1987).

Further evidence for the combined cross-linguistic influence in L3 phonological acquisition was provided by some studies by the present author (Wrembel 2011a, 2015) as well as Wrembel (2012a, b); however, they are discussed in more detail in the respective sections devoted to the overview of research on voice onset time VOT (Section 3.2.7) and foreign accentedness (Section 3.2.8).

One of the questions that remains to be investigated is whether such combined transfer involves various language sources simultaneously or successively. On the one hand, Hammarberg and Hammarberg (1993, 2005) found evidence for successive transfer (first from the L2, then as L3 proficiency advanced, from the L1), but, on the other hand, there are also observations of simultaneous transfer as reported e.g. by Barkley (2010).

3.2.6. Recent investigations into L3 phonology

The present section is the largest in scope as it covers the most recent studies into the realm of the phonological acquisition of a third language. An important landmark in the development of empirical investigations into the acquisition of L3 phonology was the publication of a Special Issue of the International Journal of Multilingualism (IJM) guest-edited by Wrembel, Gut and Mehlhorn (2010), based on the satellite workshop to ICPhS 2007 devoted to transfer in L3 phonology. The issue constituted the then state-of-the-art overview of L3 theories and research methodology, and addressed the issue of native vs. non-native cross-linguistic influence and factors that determine sources of transfer. The contributions to this special issue, including Marx and Mehlhorn (2010), Gut (2010), Llama et al. (2010), Missaglia (2010) and Wrembel (2010), were aimed at exploring the possible positive effects of previously learned languages, the application of learning strategies as well as enhanced meta-linguistic competence in multilinguals, among other issues.

The article by Marx and Mehlhorn (2010) discusses current theoretical conceptualisations of phonological third language acquisition and the role played by the second language in this process. The novelty of this paper lies in its focus on how the phonetic similarities between English and German offer possibilities for positive transfer for those learning these Germanic languages as their L2 and L3. Moreover, the authors investigated the question of multilingual learners' cognitive advantage by drawing attention to an increased level of metalinguistic awareness, a larger repertoire of phonetic-phonological parameters and a developed phonological knowledge, all of which are claimed to facilitate pronunciation learning in subsequent foreign languages. Marx and Mehlhorn provided also specific didactic suggestions for learning L3 German pronunciation based on positive transfer from L2 English.

In her contribution, Gut (2010) presented the case studies of four trilingual speakers with different first languages and English or German as their second or third languages. The study focused on vowel reduction and speech rhythm in the participants' L3 and investigated sources of CLI. The findings revealed that the multilingual subjects differed considerably in the phonological processes under investigation from the performance of native speakers. The author found conflicting evidence for L2-L3 cross-linguistic influence, with some indication of a positive effect of the L2 on the L3. There was no conclusive evidence of L1 interference in the subjects' prosodic patterns in the second and third languages. Gut (2010) concluded that the effects of the phonological properties of the L3 seem to be stronger than the cross-linguistic influence. Further, she suggested that high L2 proficiency may be a prerequisite for the L2 status to override L1 influence.

In turn, Missaglia (2010) investigated a case of triple language contact in which consecutive bilingualism constituted the basis of L3 acquisition. The articulations of vowels by child Italian-German bilinguals acquiring L3 English were analysed acoustically. It appeared that the subjects were able to distinguish between corresponding phonemes that belonged to different vowel inventories with respect to both vowel height and backness. The findings revealed that L2 German experience was beneficial for the process of the phonological acquisition of L3 English by Italian pre-puberty subjects. Missaglia put forward a hypothesis that it was not the intrinsic phonetic similarities between German and English that favoured phonetic acquisition of L3 English, but rather the bilingual children's extensive experience with the prosodic and rhythmic characteristics of both languages. This bilingual experience is claimed to have led to a correct language-specific acquisition of the segmental and suprasegmental features in L3 English.

In the remaining two contributions to this special issue, Llama et al. (2010) addressed the question of whether the 'L2 status' or the typology is the stronger predictor of cross-linguistic influence in the production of voiceless stops in stressed onset position in L3 Spanish, whereas Wrembel (2010) pursued a similar goal with the application of accentedness ratings that measured perceived foreign accent in L3 English. These papers are discussed in more detail in the sections devoted to research on specific phonetic features (i.e. Llama et al. 2010 in section 3.2.7 and Wrembel 2010 in section 3.2.8).

As far as more formal linguistic approaches to L3 phonological acquisition are concerned, the study by Cabrelli Amaro and Rothman (2010) provided new insights into the debate on the mental representations of non-native phonological systems. The authors put forward the Phonological Permeability Hypothesis in an attempt to explore differences between the acquisition of phonology in the pre- vs. post-critical period. The study focused on regressive interference between the native and non-native phonological systems, the so called cross-linguistic permeability. Based on the preliminary data from the comparison of simultaneous vs. successive English-Spanish bilinguals learning Brazilian Portuguese as their L3, Cabrelli Amaro and Rothman (2010) confirmed the predictions of their phonological permeability hypothesis – that under an influence of an L3, phonological attrition of the second language is more pervasive when the L2 acquisition happens after rather than before the critical period.

The Phonological Permeability Hypothesis was further tested by Cabrelli Amaro (2013, 2014) to establish to what extent early-acquired versus late-acquired (i.e., adult) phonological systems resist influence from a third language (L3). To this end, the author investigated the potential effects of L3 Brazilian Portuguese (BP) word-final vowel reduction on word-final vowels in L2 Spanish among sequential English/Spanish bilinguals learning L3 BP, whose acquisition of Spanish occurred either early or late in life. The study involved both perception and production, in the form of a forced-choice naturalness preference task for the former, and a delayed repetition task for the latter. In the perceptual task, the participants were to select between a token with a word-final reduced vowel (BP-like) and one with a fully-realized vowel (Spanish-like), and their accuracy scores as well as the reaction times were measured. In the production test, delayed repetitions of nonce words of the CV.CV structure were analysed for F2-F1, F1-F0, as well as differences in duration and intensity between tonic and atonic vowels. Cabrelli Amaro did not find any statistically significant differences in the perception or production data either for L3 BP or L2 Spanish irrespective of the group. The findings failed to provide support for the earlier proposed hypothesis for a differential stability of the L2 Spanish phonological system, since the stability proved independent of both the age of acquisition and the BP proficiency.

Multilingual acquisition of phonology in children has rarely been explored to date. In a series of studies Kopečková (2013, 2015) investigated segmental perception and production in young learners with diverse language combinations. The earlier investigation involved native German learners of L3 Spanish with previous knowledge of English as an L2, who were tested on their production of rhotic sounds in Spanish on two testing times, i.e. at the initial stages of L3 learning and seven months later. The study deserves special attention as it constitutes a rare example of longitudinal research into perceptual aspects of third language acquisition and involves child learners, albeit limited to a small pool of participants. In her 2015 study Kopečková examined a group of multilingual Polish children residing in Ireland, who acquired English either as their L2 or L3 alongside French or German. Having been closely matched on a range of background variables, the children were tested for their perception of cross-language phonetic similarity with respect to selected English and Polish vowel sounds. The L3 learners were found to perform with greater perceptual sensitivity than their L2 counterparts both in terms of segments similar to and different from their L1. The author tried to tease apart whether this perceptual faculty could be related to general enhanced cognitive flexibility or to a specific L2 learning experience, pointing to a greater impact of the latter. Kopečková emphasised that perceptual processes in both L2 and L3 language acquisition relied on the basic cognitive mechanism of equivalence classification; however, L3 child learners seemed "to be in a better position to attune their perceptual processing to L3 speech, i.e. to re-sensitise and modify their perception of L3 sounds in the direction of the sound categories established during L1 and L2 acquisition" (2015: 86). The author concluded that young learners were thus able to capitalise on their extended experience with additional sounds systems.

Another comprehensive study was conducted by Sypiańska (2013) who analysed the multilingual acquisition of vowels by trilingual speakers of L1 Polish, L2 Danish and L3 English. Selected vowels including Polish /ɛ/, Danish /e, ɛ, æ/ and English /e/ were analysed acoustically in terms of the first two formants (F1 and F2), then compared cross-linguistically and against monolingual baseline data. The analysis demonstrated a general tendency for the multilinguals' vowel space to be subject to reshaping in all three languages and to become less peripheral. The participants' L3 English vowel formants were found to exert a significant influence on both the F1 and F2 of their L1 Polish and L2 Danish vowels. Further, the vocalic qualities in all three languages of the multilingual speakers' linguistic repertoire were shown to vary from the monolingual baseline data. Based on her findings, Sypiańska (2013, 2014) postulated the existence of a global language entity in multilingual speakers, i.e. an entity composed of multilingual's component languages which exhibit qualitative phonetic differences from the same languages in monolingual speakers.

Some of the most recent investigations in the area under study were presented at the SLE 2014 workshop on Advances in the Investigation of L3 Phonological Acquisition convened by Jennifer Cabrelli Amaro and Magdalena Wrembel. The preliminary findings of these studies by Onishi (2014), Llama and López-Morelos (2014), Gabriel et al. (2014), Lechner and Kohlberger (2014) shall be discussed below.

The effect of L2 experience in L3 perception was examined by Onishi (2014). Perception tasks involving a forced-choice identification and AXB discrimination of various minimal contrasts in Japanese and in English were administered to a group of multilingual speakers with L1 Korean, L2 English who have been learning Japanese as their L3 and to monolingual English controls. In order to determine whether the perceptual ability in L2 English had any influence on the perception of L3 Japanese contrasts, the author investigated potential correlations between the participants' performance on different experimental tasks. The analysis demonstrated that there was a positive correlation between English minimal pairs identification/discrimination by the Korean participants and their performance on several, though not all, Japanese contrasts. The results suggest that learners can refer to the phonological categories of all of their background languages when identifying L3 categories. However, some of the reported perceptual difficulties may indicate an increased general sensitivity to speech sounds in L3 learners rather than an explicit use of L2 categories. This enhanced auditory awareness is ascribed by Onishi (2014) to previous L2 learning experience.

In turn, Llama and López-Morelos (2014) investigated the production of voiceless stops by trilingual children and teenagers who were heritage speakers of Spanish with French and English as their second or third languages. The study aimed to analyse to what extent the participants were able to produce target-like VOT values in all their languages on the basis of their reading of three word lists. The results of the acoustic measurements demonstrated high variability, especially in VOT productions by trilingual children, therefore, the authors were not able to make any strong claims about how distinct their VOT systems were in the respective languages. Llama and López-Morelos tried to account for their inconclusive findings with regard to such factors as age, language dominance cross-linguistic influence and to interpret them in relation to current transfer models proposed for multilingualism.

Gabriel et al. (2014) investigated the production and perception of a non-native accent in L3 French in multilingual learners with Mandarin Chinese as a heritage language (HL). The aim was to determine whether multilingual learners demonstrate an advantage over monolinguals with respect to selected segmental and suprasegmental aspects of pronunciation. To this end, the study compared the performance on speech rhythm, voice onset time, and global foreign accent in French learnt as a foreign language in three learner groups including: multilingual German/HL Chinese, monolingual German and monolingual Chinese speakers. The data collection procedure involved reading a list of sentences in French, German and Chinese that were subject to subsequent analysis. The results of the VOT analysis suggest that knowing Chinese in the case of multilingual learners was neither an advantage nor a disadvantage because the subjects developed compromise VOT values intermediate between German and Chinese. The authors reported mostly preliminary results of their study pointing to multilinguals' advantage over German monolinguals with respect to speech rhythm whereas monolingual Chinese learners could rely on their L1 in the production of a French speech rhythm.

Lechner and Kohlberger (2014) examined phonetic transfer in vowel production in L3 English in bilingual children and adults. The participants included subtractive successive bilinguals of diverse L1 backgrounds (Turkish, Russian, Persian) with German as an L2 acquiring English as an L3, as well as a monolingual German control group with L2 English. The data collection involved reading a passage, story telling and free speech in all three languages, and the analysis focused on vowel formants (F1 and F2) and durations. Preliminary results reported by the authors indicated some advantage of the multilinguals in their phonetic production over their monolingual peers, whereas transfer effects were observed both from the speakers' L1 and L2.

Summing up, the trends observed in the survey of the most recent studies point to a broader scope of investigation encompassing both perceptual and production experiments, a wider range of language repertoires (e.g. including non-Indo-European languages), and a more expanded age range of the participants, i.e. allowing for multilingual child acquirers.

The following subsections concentrate on a selected range of features that constitute the focus of the present contribution. Consequently, the following survey sections cover four strands of research, including: studies into voice onset time acquisition in L3 (section 3.2.7), foreign accentedness ratings in L3 (section 3.2.8), studies on metaphonological awareness in multilinguals (section 3.2.9) as well as investigations into interdependency between the above mentioned features (section 3.2.10).

3.2.7. VOT studies

Research into the acquisition of foreign language phonology has focused on specific phonetic or phonological features, frequently including the voice onset time (VOT) of initial stop consonants. VOT is commonly acknowledged as a feature correlated with the degree of perceived foreign accent. Further, the precise nature of acoustic measurements of VOT enables accurate statistical analyses and, consequently, a verification of research hypotheses.

The acquisition of different voice onset time (VOT) patterns across languages has been widely investigated in the field of Second Language Acquisition (SLA). An overview of SLA literature provides numerous evidence of the transfer of VOT values from the first to the second language, especially at the lower levels of L2 proficiency (e.g. Flege 1987; Flege and Hillenbrand 1987). Flege and associates tried to explain this inability to distinguish between differently aspirated plosives in L1 vs. L2 by referring to the mechanism of equivalence classification, which blocks the formation of a new phonetic category. Further, more advanced learners were also commonly reported to create hybrid VOT values intermediate between their L1 and the target L2 (Flege 1991, Gurski, 2006). However, only the most proficient L2 users were found to approximate native-like VOT durations (Flege 1987). A more detailed discussion of theoretical accounts of various VOT acquisition patterns as well as potential influencing factors is presented in Wrembel (2015).

Notably, the phenomenon of voice onset time has also began to receive attention in Third Language Acquisition and VOT values have been used to determine cross-linguistic influence in the L3 phonological system. Tremblay (2007) was the first to investigate VOT acquisition patterns from a multilingual perspective. The participants included four L1 English/L2 French bilinguals at the early stages of the acquisition of L3 Japanese. VOT measurements demonstrated comparable durations for L2 French and L3 Japanese which were much shorter than those in L1 English. The author interpreted the results as an indication of the L2 status on L3 phonological acquisition, although the L3 VOT values approximated not only L2 French but also the native Japanese target norms. No task effect was reported in the study as L3 VOT patterns were not significantly different in the word reading vs. delayed repetition task. An obvious limitation of the study was a very limited number of participants, which did not allow to draw far reaching generalisations.

In turn, a study by Llama et al. (2010) was based on a mirror-design methodology involving two groups of L3 Spanish learners with either L1 English/L2 French or L1 French/L2 English. Such a design was aimed at testing the research hypothesis whether the L2 status or language typology acts as a factor conditioning VOT patterns in the third language. The experiment consisted in reading lists of target words with onset voiceless plosives in the subjects' L2 and L3. The results demonstrated that CLI from the L2 rather than typological proximity or L1 transfer alone turned out to be a stronger predictor for the L3 VOT measures. Nonetheless, there were also some indications of the interaction of native and non-native influences on L3. A shortcoming of this study is the lack of VOT acoustic measurements in the participants' L1s and the reliance on the literature reference values as a baseline instead.

Wunder (2010), as mentioned in an earlier section, analysed text reading samples of eight L1 German speakers with respect to the VOT measures in their L2 English and L3 Spanish. She reported mixed results pointing to either an L1 effect or a combined L1 German and L2 English cross-linguistic influence on VOT patterns in L3 Spanish. The majority of the acoustic measurements were categorised as 'hybrid' values as in these cases it was not possible to determine whether the sources of CLI were the L1 German or native Spanish values. Wunder interpreted her results as failing to provide support for Hammarberg and Hammarberg's (2005) findings on an L2 status effect in L3 acquisition of phonology.

Somewhat different patterns were observed by Sypiańska (2013) who examined the voice onset time of word-initial voiceless plosives in multilingual speakers of L1 Polish, L2 Danish and L3 English. The author reported a combined CLI from the first and second language on VOT patterns in L3 English as well as a regressive transfer from the L3 reflected in the prolonged VOT values in L1 Polish and L2 Danish in the case of the trilingual group as compared to a bilingual control group. Sypiańska concluded that her findings point to the phenomenon of what she referred to as a 'global language entity' attested in the mutual influence and interaction between all the component languages of multilingual speakers.

An investigation into the trilingual acquisition of VOT patterns was also undertaken by the present author in order to broaden the research perspective provided by perceptual rating studies (cf. Wrembel 2012a, 2012b) by focusing on laryngeal durations as a selected dimension of foreign accentedness.

The first exploratory study (Wrembel 2011a) demonstrated that the multilingual subjects contrasted between VOT length in their three phonological systems, i.e. L1 Polish, L2 English and L3 French. Notably, they produced voiceless plosives in L2 English with target-like long-lag values, whereas their implementations of /p, t, k/ in L3 French had significantly longer VOT values than those of French monolinguals. The reported VOT durations in L3 French were found to be intermediate between the L1 Polish and L2 English mean VOT and as such the hybrid values corroborated the co-existence of the L2 effect and underlying L1 interference, thus substantiating the assumption of a combined cross-linguistic influence on L3 acquisition (see Section 3.2.7).

A related investigation (Wrembel 2015) intended to compare studies on VOT patterns in various acquisition settings, i.e. when a foreign language is acquired as a second vs. third/additional language or as a new variety. The language repertoire of the study group involved multilingual subjects with L1 German, L2 English and L3 French. Similarly to Wrembel (2011a), the mean VOT values in L3 French were reported to exceed considerably the target native French durations, yet they were lower than the respective values in L1 German and L2 English. However, due to nearly identical patterns of distribution displayed by the subjects' L1 and L2 VOT, it was not possible to tease apart the influence from the native vs. non-native languages on shaping L3 acquisition patterns. Nonetheless, the compromise VOT values in L3 were interpreted as evidence for the co-existence of L1 and L2 interference. In order to address the question of whether there are fundamental differences in the process and outcome of phonological acquisition as a function of various settings, the present author suggested that the category assimilation observed in the acquisition of L3 VOT seems to be of a different nature than the one reported in the SLA literature. While research on the acquisition of L2 phonology demonstrated a prevailing tendency for a hybrid VOT between the native and target values (e.g. Flege 1987, Waniek-Klimczak 2011), studies on new varieties of English reported a strong influence of L1 VOT durations (e.g. Shahidi and Rahim 2011, Wissing 2005). In turn, L3 phonology has been mostly shown to draw on both the native and non-native language systems of the multilingual subjects, specifically when referring to VOT acquisition patterns.

3.2.8. Accentedness ratings in L3

Ratings of perceived global foreign accent have been widely applied in second language acquisition (SLA) research (e.g. Flege 1988; Gallardo del Puerto Gómez Lacabex and García Lecumberrri 2007; Piske et al. 2001); nonetheless, so far they have not been used extensively in studies on L3 phonology. Hammarberg and Hammarberg's (1993, 2005) seminal case study was the first notable exception and its results and implications are discussed in more detail in section 3.2.4.

A preliminary series of foreign accentedness studies were performed by the present author to explore the sources of cross-linguistic influence in third language phonological acquisition, while, at the same time, to control for the confounding variable of language typological proximity (Wrembel 2010, 2012a, b). The first exploratory investigation tested the perception of global foreign accentedness in L3 and the ability to identify the speakers' mother tongue based on the L3 speech samples (Wrembel 2010). The results fully substantiated Hammarberg's findings on L2-accented speech being dominant at the early stages of L3 phonological acquisition and are discussed in more detail in the subsection on L2 status effect (see section 3.2.4).

Two follow-up studies employed a mirror-design methodology and involved the following language combinations: L1 Polish, L2 English, L3 French (Wrembel 2012b) vs. L1 Polish, L2 French, L3 English (Wrembel 2012a). The results did not provide support for previous findings on the prevalence of the L2 status in L3 phonological acquisition (e.g. Hammarberg and Hammarberg 2005, Wrembel 2010). However, based on the patterns of L1 identification, both studies identified the native language as the main source of CLI, demonstrating also a partial effect of L2 phonology. Cross-linguistic influence from the subjects' L1 Polish dominated in their phonetic performance in the third language irrespective of language proficiency or the stage of L3 phonological acquisition. In an attempt to avoid any typological bias, the multilingual repertoires involved different language families (i.e. Slavic, Romance and Germanic). However, it proved difficult to ensure complete typological neutrality, with French and English being more typologically similar in terms of lexis and syntax than Polish and English or Polish and French. Nonetheless, the languages involved were more distinct in terms of phonology, particularly with respect to prosodic and temporal patterns as well as the respective phonemic structures.

An interesting study on the perception of foreign-accented Polish was conducted by Szpyra-Kozłowska and Radomski (2012). Although the authors did not explicitly place their investigation in the third language acquisition perspective, it may be assumed that Polish was an additional foreign language (i.e. L3 rather than L2), given the participants' different L1 backgrounds including American, British, French, Spanish, Chinese and Russian. Through the application of accentedness ratings, the study aimed to explore the listeners' comprehensibility, ability to identify the foreign accent as well as their attitudes to foreign-accented Polish. A note-worthy novelty of this contribution were impressionistic evaluations of various accents provided by the raters.

3.2.9. Studies on metaphonological awareness

Metalinguistic awareness in the area of phonology has not been investigated extensively within the realm of second language acquisition, let alone from the multilingual perspective. Therefore, the present overview of research in this field will start with some investigations related to the L2 acquisition setting (Kennedy and Trofimovich 2010, Mora et al. 2014, Venkatagiri and Levis 2007, Wrembel 2005, 2011) and further expand to L3 phonological acquisition (Wrembel 2013, 2015).

Within the SLA perspective, Venkatagiri and Levis (2007) investigated the interrelation between phonological awareness and speech comprehensibility. To this end, a large battery of tests was administered, including 14 tests of phonological awareness (i.e. phonological blending, segmentation, manipulation, spoonerism, rhyming, alliteration) and 3 tests of phonological short term memory (i.e. non-word recall). The comprehensibility ratings performed by native speaker raters showed a positive correlation between composite phonological awareness scores and rated comprehensibility, as well as between phonological awareness and phonological short term memory. The findings demonstrated that learners with superior explicit phonological knowledge were perceived to be more intelligible foreign language speakers.

In turn, Kennedy and Trofimovich (2010) carried out a classroom study in which they examined how language awareness is related to the quality of L2 pronunciation assessed by means of accentedness, comprehensibility and fluency ratings. Language awareness was classified as either quantitative or qualitative based on journal entries made by the learners. The analysis indicated that there was a relationship between L2 pronunciation ratings and the number of qualitative (but not quantitative) language awareness comments, i.e., higher pronunciation ratings were associated with a greater number of qualitative language awareness comments, related to using pronunciation to convey the intended message.

In a recent investigation Mora et al. (2014) aimed to assess the phonological awareness of Spanish learners of English by adapting a foreign-accent-mimicking task to discern learners' implicit knowledge about non-distinctive phonetic differences between VOT durations in their L1 Spanish and L2 English. The authors stated that investigations into phonological awareness should involve implicit knowledge about the phonological system of the target language as well as cross-linguistic phonetic differences. The participants were found to be able to modify their native short-lag laryngeal timing patterns by producing longer VOT values in L2 English as well as in their imitations of English-accented Spanish words. The authors interpreted their results as evidence for an implicit phonological awareness of cross-linguistic differences in VOT realizations developed by Spanish learners of L2 English.

Also the present author has conducted a number of investigations into phonological awareness from various perspectives. In the first longitudinal study of this kind, Wrembel (2005) explored the impact of analysed phonological knowledge and metalinguistic awareness on pronunciation performance in English learnt as a second language. It was found that the participants in the experimental group, who were equipped with declarative phonological knowledge, outperformed the controls, who relied solely on procedural knowledge developed through practical pronunciation training in L2 English provided for both groups. The findings substantiated the claim that explicit phonological knowledge and awareness raising contribute to the development of L2 phonological competence and may be considered as predictors of pronunciation attainment in a foreign language. Further, the foreign-accent mimicking task proved to be a valuable method of assessing phonological awareness.

A related study (Wrembel 2011) was intended to complement previous findings by tapping into the learners' self-perception of metaphonological awareness. To this end, spontaneous L2 speech productions were investigated by means of verbal protocols to elicit the participants' self-reflection and pronunciation monitoring strategies. The results pointed to a considerable degree of analysed knowledge reflected in the explicit references to English phonology, instances of self-repair and self-reflection on L2 pronunciation learning strategies. Nevertheless, the level of control appeared to be not so advanced as the participants exhibited problems with monitoring phonetic form during their L2 oral performance.

Further studies on various aspects of metaphonological awareness have gone beyond the SLA perspective and expanded the scope of investigation to a multilingual acquisition setting. Wrembel's (2013) piloting study with the application of stimulated recall protocols investigated metaphonological awareness of Polish speakers with advanced proficiency in L2 English acquiring French or German as their third languages. The major aim was to explore the cognitive aspects of attention and noticing in L3 oral production and to investigate if trilingual speakers of typologically related vs. unrelated sets of languages resort to their first or second language in L3 speech and whether they were aware of this phenomenon.

To ensure wider cross-linguistic comparisons, a similar design was replicated in another study (Wrembel 2015) that focused on a different multilingual repertoire (i.e. L1 German, L2 English, L3 Polish, and other Slavic Ln). The research design involved quasi-concurrent retrospective and introspective protocols, in which the participants were asked to modify and comment on their phonological performance in L3 Polish after listening to excerpts of their text reading recording. The analysis of metaphonological awareness was both qualitative and quantitative in nature, and identified various manifestations thereof such as L3 self-repair, corrections of mispronunciations, self-awareness of problems in L3 pronunciation or reflective comments on the process of pronunciation learning. It was shown that in the case of learners acquiring their third or additional languages, metaphonological awareness constitutes an important component of multilingual competence as well as that it entails an interaction of metalinguistic awareness and cross-linguistic awareness, as indicated in the literature (cf. Jessner 2006).

3.2.10. Interdependency studies

Relatively few studies to date have attempted to investigate the correlation between at least two of the selected focal features, e.g. between voice onset time and accentedness ratings or between accentedness and metalinguistic awareness, not to mention an interdependency research allowing for all these properties. However limited, this research has been conducted solely in the second language acquisition setting, yet it will be included in this survey to form a foundation for the present empirical contribution embedded in the multilingual acquisition perspective.

Three of the studies reviewed in the subsection on metalinguistic awareness (i.e. section 3.2.9), including Kennedy and Trofimovich (2010), Mora et al. (2014), Venkatagiri and Levis (2007) can be categorized as interdependency research. While focusing on different measures of phonological awareness in a foreign language (i.e. literacy related tests vs. the analysis of diary comments vs. a foreign-accent-mimicry paradigm respectively), the afore-mentioned studies investigated its relation with perceived L2 pronunciation performance. The results indicated that there is some correlation between phonological awareness and the global ratings of comprehensibility (Venkatagiri and Levis 2007) as well as accentedness, comprehensibility and fluency ratings (Kennedy and Trofimovich 2010).

Mora et al. (2014), in turn, operationalised phonological awareness as the participants' implicit ability to discern cross-linguistic differences in VOT realizations in L1 Spanish, L2 English and English-accented Spanish. Moreover, they explored the interdependency between voice onset time and accentedness ratings. It was found that VOT measures of L2 English word tokens were strongly correlated to accentedness ratings by English native speakers, i.e. the subjects with longer VOT values were evaluated perceptually to be less foreign accented in their L2 English. On the other hand, the correlation between VOT durations in Spanish words imitating English accent and the accentedness rating of these words performed by Spanish native speakers were much weaker and fell below the significance level. These results yielded partial support for the intercorrelation between the developed implicit phonological awareness, as assessed through VOT accuracy measures, and accentedness ratings.

As far as the interdependence between VOT and the perception of foreign accentedness is concerned, there have been a limited number studies following this line of investigation. The majority of early SLA studies found correlations between VOT and global accent for L2 learners (i.e. Flege, 1984; Flege and Eefting, 1987, Major 1987) and their findings have been mostly confirmed by more recent research (e.g. Jilka, 2000; Neuhauser, 2011; Riney and Takagi 1999) with the exception of Lein et al.'s (2015). Data collection procedures were based mostly on text or sentence reading, thus generating fairly controlled data.

Major (1987) and Flege and Eefting (1987) were the first to establish a relation between global accentedness and voice onset time in the context of second language acquisition. Specifically, Flege and Eefting (1987) examined how voiceless plosives /p, t, k/ were implemented as aspirated stops in English, but as unaspirated stops in Dutch by testing their identification and production by participants with L1 Dutch and L2 English. The subjects were found to be able to distinguish between Dutch and English /t/ and to produce a longer mean VOT in L2 English than L1 Dutch /t/. However, this ability was significantly greater for proficient than non-proficient subjects, and it proved proficiency to be a function of global accentedness. Moreover, the authors demonstrated that highly competent Dutch speakers of L2 English formed a new category for English /t/, yet they overshot their English VOT values compared to monolingual reference norms, dissimilating between VOT values in the two languages. In his Speech Learning Model Flege (1995) accounted for this finding putting forward a hypothesis that L2 learners tend to maintain a contrast between phonetic categories in a common L1-L2 phonological space (1995: 239).

Riney and Takagi's (1999) study on the correlation between global foreign accent and VOT among Japanese EFL speakers followed closely the previous findings and provided further support for Major's (1987) claim that linked global and discrete measures of accentedness. Interestingly, VOT measures of voiceless plosives in English and Japanese taken at two testing times separated by 2 years did not show changes over time, as demonstrated by Riney and Takagi (1999).

A different pattern was reported in one of the most recent investigations of this kind by Lein et al.'s study (2015) into the interaction between voice onset time and global foreign accent in German-French simultaneous bilinguals. In this study, VOT durations of the voiceless plosive /k/ were elicited through a spontaneous speaking task in French and German and further correlated with accentedness ratings. The results indicated that the bilingual participants exhibited distinct VOT mean values in the respective languages (i.e. higher in German and lower in French), which is in line with the assumption of language separation and Flege's (1995) claim on the dissimilation of L1/L2 VOT values. Nevertheless, no systematic relationship was found between foreign accent and VOT since there were speakers with a native-like accent and a deviant VOT, as well as speakers with a non-native accent in one language but distinct VOT categories in each language. Although the findings were inconsistent with previous studies by Flege and Major, Lein et al. (2015) tried to argue that target-like VOTs do not preclude an accent in other phonological domains as demonstrated in the research on heritage speakers whose VOT was in the native-like range, yet who displayed perceivable foreign accents (cf. Oh et al., 2003). The authors emphasise that the results proved rather unexpected and partially inconsistent with their earlier findings on the speakers' global accent (Kupisch et al., 2014). Further, they put forward an explanation that since VOT is a highly variable phenomenon it may not function as a major determinant in the perception of global accent, although such an account is inconsistent with the majority of previous research findings.

All the interdependency studies surveyed in this subsection have explored the relationship between the selected focal features in the second language acquisition context. The current series of interconnected studies reported in the empirical part aims to expand the research perspective to explore possible relations between voice onset time, accentedness ratings and metaphonological awareness from the third language acquisition perspective. Therefore, the present contribution is innovative from the point of view of the scope of investigation as well as the adopted acquisition setting.

3.3. Methodological considerations

Several methodological issues are involved in research on L3 phonology due to the complexity of the process of multilingual acquisition, the presence of more than two language systems and various conditioning factors (cf. Cabrelli Amaro 2012). These considerations will be discussed in this section including the scope of investigation (i.e. perceptual vs. productive studies), key research areas (i.e. the selection of phonetic properties), language groupings, as well as proficiency measures.

As for the scope of investigation in third language phonology, the majority of the studies have focused on the production patterns in L3 phonology, with very few investigating L3 speech perception. This imbalance reflects the trends reported in SLA phonological research (cf. Gut 2009); however, in the case of third language acquisition perceptual studies are even more underrepresented. As pointed out by Cabrelli Amaro (2012: 50) the most influential models of L2 speech have used perception to explain the acquisition of second language phonology, therefore, it would be beneficial to test the production-perception interface from the point of view of L3 phonology as well. Nonetheless, research into perceptual aspects of L3 phonological acquisition has been very limited so far with the exception of studies by Kopečková (2013, 2015), Cabrelli Amaro (2013) and Onishi (2014), which were reviewed previously (see section 3.2.6).

As far as key research areas in L3 phonological acquisition are concerned, there are a few selected phonetic and phonological features that have attracted most of scholars' attention. Similarly to the research trends reported in SLA (cf. Gut 2009), segmental properties have received far more coverage than suprasegmentals. The majority of investigations into L3 phonology have concerned segmental properties such as voice onset time or vowel formants.

For the reasons related to the precise nature of the acoustic measurements of VOT, a number of studies focused on the acquisition of voice onset time patterns, including Tremblay (2007), Llama et al. (2010), Wunder (2010), Wrembel (2011a, 2015), Sypiańska (2013), and Llama and López-Morelos (2014); see section 2.3.7 for a detailed discussion thereof. Alongside stop consonants, vowels constitute one of the most widely studied classes of sounds in SLA research, as indicated by Hansen Edwards and Zampini (2008). Parallel trends are evidenced in L3 acquisition research with vocalic sounds being probably the second most investigated class of features, as reflected in studies by Missaglia (2010), Sypiańska (2013), Lechner and Kohlberger (2014) and Kopečková (2015).

Suprasegmental features have not been investigated widely yet, with the notable exceptions of Louriz's (2007) examination of word stress; Gut's (2010) case study into speech rhythm and vowel neutralization; Cabrelli Amaro (2013) vowel reduction; or Gabriel et al.'s (2014) preliminary investigation of speech rhythm. Furthermore, research into the re-setting of the articulatory basis in the third language initiated by Hammerberg and Hammerberg (1993, 2005) and continued by Wrembel (2007) as well as the present author's research on global foreign accentedness in L3 (Wrembel 2010, 2012a, 2012b) have reached beyond the traditional segmental perspective. Nonetheless, in order to gain a more holistic view of the field, more research on prosody-related aspects of third language acquisition would be particularly welcome.

With respect to the language groupings that have constituted the subject of investigations in L3 studies, their scope has been also rather limited to date. The majority of the studies focus on Indo-European languages with a prevalence of Germanic and Romance languages as parts of the subjects' multilingual repertoires. Some of the investigations included Slavic languages, either as the L1 or L3 (cf. Marx and Mehlhorn 2010; Kopečková 2015; Wrembel 2010, 2011a, 2012a, 2012b, 2015). Only a few of the studies involved non-Indo-European languages as part of the participants' multilingual repertoire, thus providing more typologically distant sets, e.g. Japanese in Trembley (2007) and Kamiyama (2007), Arabic in Benrabah (1991), Mandarin Chinese in Gabriel et al. (2014), and Korean and Japanese in Onishi (2014). A call for international co-operation between L3 scholars was issued by Cabrelli Amaro (2012: 53), who stressed the necessity of increasing sample sizes and creating language corpora consisting of more diverse language pairings. She further raised the question of what constitutes a typological relationship between languages; whether it relates to the linguistic system as a whole, the phonological system, or rather a selected phonetic/phonological property. Determining the typological relationship appears to be an important methodological consideration.

Another debatable methodological issue is related to the assessment of language proficiency. Cabrelli Amaro (2012: 52) advocates that "objective phonological proficiency testing must be implemented and used in conjunction with a measurement of global proficiency to classify participants". Standard proficiency measures should ensure the validity of a study; however, to date a range of various assessment procedures is applied including self-rating, language class membership or vocabulary size tests. It would thus be necessary to establish some methodological norms and to design more streamlined methods for proficiency assessment.

Finally, the use of an appropriate research design has also been subject to a methodological debate. To date, a vast majority of studies on third language phonology has been cross-sectional in nature, as such studies are logistically much more convenient to conduct. We have been faced with a dearth of longitudinal investigations, with Hammarberg and Hammarberg's (1993, 2005) seminal case study of Sarah Williams as a notable exception. Some studies have at least attempted more than one data collection time (e.g. Kopečková 2013); however, longitudinal research still remains a much desired design that might provide valuable insights into developmental patterns of third language phonological acquisition from its initial stages.

3.4. Conclusion

This chapter aimed to provide an in-depth overview of the literature of the field, ranging from the earliest accounts of multilingual case studies to the most recent empirical investigations into the acquisition of L3 phonology. To sum up, the existing body of research points to the complex and sometimes conflicting patterns of transfer from native and non-native languages; however, it seems still insufficient for providing a full account of the sources and directions of the cross-linguistic interference, as well as its conditioning factors. Thus the present contribution aims to bridge the gap in the literature by providing more comprehensive insights into the intricacies involved in the process of the phonological acquisition of a third language.

Chapter 4

Study I – Accentedness ratings in L3

4.1. Introduction to empirical part

This section is intended as an introduction to the empirical part of the work, encompassing a series of three independent studies conducted by the present author on four groups of participants. Firstly, the overall goals of the undertaken empirical investigations will be presented, followed by an outline of the selected methodological approaches, and a discussion of the participants' profiles.

4.1.1. Three studies, three approaches

The major goal of the series of studies reported in the present dissertation was to gain a more comprehensive understanding of L3 phonological acquisition. To this end, a holistic approach was pursued combining different methodologies of data collection and analysis. Consequently, three studies were designed and conducted in parallel on the same groups of participants, featuring:

Study 1: Accentedness ratings
Study 2: Voice onset time (VOT) acoustic measures
Study 3: Metaphonological awareness protocols.

The respective studies corresponded to three selected measures of phonetic performance in the third language, namely:

1) perceived global accent in L3
2) VOT as a correlate of foreign accentedness
3) metalinguistic awareness in L3 pronunciation

The three chosen different approaches to data collection and analysis stemmed from the original assumptions of the studies and resulted in the following designs:

Study 1: an online rating questionnaire including four parameters of foreign accentedness, comprehensibility, pronunciation correctness and L1 identification; with the ratings performed on L3 speech samples.

Study 2: acoustic measurements of voice onset time in all three languages of the participants (i.e. L1, L2 and L3) elicited through controlled reading tasks. VOT measurements compared across languages and with control reference values.

Study 3: introspective and retrospective oral protocols aimed at investigating metaphonological awareness. A complex codification system for a qualitative and quantitative analysis of generated data and a composite measure of metaphonological awareness was proposed.

A battery of instruments was designed for the purpose of the studies, including:

– an online accentedness rating questionnaire – Study 1
– a biodata questionnaire for the raters (as part of the online accentedness rating questionnaire) – Study 1
– a language learning biography questionnaire (all studies)
– 4 lists of target words in carrier phrases for VOT elicitation (in the form of ppt) – Study 2
– reading passages in 4 languages – Study 2, 3
– semi-spontaneous questionnaire format for eliciting introspective and retrospective verbal protocols – Study 3

Data elicitation procedures resulted in the creation of a large database of recordings featuring:

– word lists in carrier phrases in participants' L1, L2 and L3
– read passages in L3
– spontaneous speech samples (data not complete)
– oral verbal protocols of metaphonological awareness
– control recordings involving groups of native speakers of English, French and German, performing word list and passage reading.

The database can be used as a source for data analysis for further studies on L3 phonological acquisition.

The present integrated investigation was preceded by a number of pilot studies by the present author aimed at testing and modifying selected methodological designs. These preliminary studies involved separate data collections. The investigated language repertoires were similar due to a limited availability of multilingual leaners with a rather homogeneous profile. The results of these exploratory studies were reported in Wrembel (2011a, 2012a, 2012b, 2015).

The main idea behind the integrated series of studies was to run the same experiments on fairly homogeneous groups of multilingual participants to enable subsequent across studies and across groups comparisons as well as intercorrelation analyses between selected phonetic/phonological features and the participants' variables. A further advantage of this design involves a large pool of subjects, i.e. a total of 128 participants with different language repertoires, which is very infrequent for phonology-oriented or other studies on multilingual populations.

Recapitulating, the present investigation has aimed to overcome methodological limitations of studies on third language phonology (cf. Cabrelli Amaro 2012 and section 3.3 for a review thereof) by incorporating a holistic approach, combined methodologies, different language repertoires, and a large subject pool.

Another important aim of this series of studies has been to test the assumptions of current theoretical models proposed for multilingual acquisition including CEM, the L2 Status Model, and TPM (see Chapter 2 for a detailed discussion). The existing models were originally proposed to account for the acquisition of the morphosyntax of a third language; however, their major tenets have been extended to other domains of language as well. Specifically, the present investigator intends to explore what factors influence phonetic performance in a third language and to what extent cross-linguistic influence is conditioned by either the L1 effect, L2 status, typological proximity or the mixture of these factors. It is hoped that the findings will contribute to a better understanding of the complex process of the acquisition of L3 phonology and possibly lay foundations for the creation of a separate theoretical model of phonological acquisition in a multilingual setting.

The three aforementioned studies will be outlined and discussed separately in detail in three consecutive chapters (4, 5 and 6) and a joint comparison will be presented in Chapter 7.

4.1.2. Language groups

The aforementioned three studies were conducted in parallel on four groups of participants with complementary language triads including Polish, English, German and French in various constellations. The criteria for group selections encompassed a relative homogeneity and comparability of multilingual group members as well as purely practical considerations, i.e. the availability of language combinations in the Polish context.

The selected groups of participants involved the following language triads:

Group A: L1 Polish, L2 English, L3 French
Group B: L1 Polish, L2 English, L3 German
Group C: L1 Polish, L2 German, L3 English
Group D: L1 Polish, L2 French, L3 English

The groups varied in terms of two dimensions, i.e. the L2/L3 language status and typological proximity. As far as the parameter of language status is concerned, the selected groups followed a mirror design, with L1 Polish being kept as a stable variable, and the foreign languages assuming either the L2 or L3 status in different language pairings, thus resulting in mirror designs for the following pairs of groups:

a) A and D (L2 English, L3 French vs. L2 French, L3 English)
b) B and C (L2 English, L3 German vs. L2 German, L3 English).

The rationale behind this design was to tease apart the status of a third language from that of a second language and analyse its potential impact on the results of the respective groups in the three studies based on comparable language triads.

The second parameter of typology controlled for the typological distance between the languages involved in each of the multilingual groupings both in the sense of the systemic, i.e. global proximity as well as more local typological relatedness related to the phonetic or phonological features under investigation. Consequently, the following typological proximity relations were recognized to hold:

a) English and German (typologically related Germanic languages)
   i stress-timed languages (global perception)
   ii aspirating languages (VOT)
b) French and Polish (Romance vs. Slavic – no evident systemic relationship)
   i syllable-timed languages (although Polish has a more mixed status)
   ii voicing languages (VOT).

It follows that in terms of general typology, the experimental groups could be subdivided into two categories involving the following language pairings:

a) typologically related
   Group B: L1 Polish, L2 English, L3 German
   Group C: L1 Polish, L2 German, L3 English
b) typologically more distant
   Group A: L1 Polish, L2 English, L3 French
   Group D: L1 Polish, L2 French, L3 English.

Across-group comparisons allowing for the above mentioned language pairings are expected to account for the role of the L2/L3 status and typology in shaping the sources and directionality of cross-linguistic influence in the acquisition of third language phonology, which is the underlying goal of this series of investigations.

4.1.3. Participants' profiles

Based on the language history questionnaire that the participants were requested to fill in at the end of the recording session, a detailed participants' profile was created for each group. The generated bio data will serve as participants' variables and will be entered into statistical analyses presented in the results sections.

4.1.3.1. Group A: L1 Polish, L2 English, L3 French

Group A comprised 39 participants, i.e. Polish university students of English philology who have been learning French as an additional foreign language (L3) (see Table 2 for the participants' profiles). There were 32 females and 7 males involved in the study and their mean age was 20.3 years (SD=1.4. range 19-25).

Their competence in L2 English was advanced, ranging from the B2 to the C1 levels according to the classification of the Common European Framework of Reference (CEFR 2011), based on their practical English exam results. The length of formal training in English (L2_YFT) was 11.8 (SD=2.1); whereas the mean age of onset of learning (L2_AOL) equalled 8.5 years (SD=2.2). The participants had been learning French as their third language (L3). i.e. a subsequent foreign language after English. Their proficiency level in L3 French ranged from A1 (elementary) to B2 (intermediate) based on the internal course placement tests. The average exposure to French (L3_YFT) equalled 3.7 years (SD=2.3) with the mean age of onset of learning (L3_AOL) being 16.3 years (SD=2.7). The self-declared evaluation of proficiency level in L3 French pronunciation was on average 2.6 (SD=0.6) on a 5 point scale.

As far as a stay in French speaking countries is concerned, 50% of the participants had paid such a visit and the average length of stay for the whole population was 1.7 weeks (SD=4.7). Further. the participants were asked to evaluate their general competence in L3 French as well as their L3 pronunciation on a 5 point scale ranging from 1=very poor to 5=very good. The self-evaluation in the former case equalled 2.4 (SD=0.9). whereas in the latter 2.6 (SD=0.8). The number of foreign languages known by the participants was on average 2.6 (SD=0.6). The participants had received general linguistic training during their course of studies of English philology and phonetic instruction in L2 English; however, no regular practical training in French pronunciation was provided during the L3 French course. The participants' profile for Group A is presented in Table 2.

Table 2. Participants' profile of Group A

N=39	L2 English				L3 French				
	Age	Prof	YFT	AOL	Prof	YFT	AOL	Eval max 5	Stay_L3 weeks
M	20.3	B2/C1	11.8	8.5	A1/B1	3.7	16.3	2.6	1.7
SD	1.4		2.1	2.2		2.3	2.7	0.6	4.7

YFT – years of formal training, AOL – age of onset of learning, Prof – proficiency (both as a proficiency level in CEFR and as assessed in a placement test (max 5), Eval – self-evaluation of L3 pronunciation, Stay – length of stay in L3 speaking countries (in weeks)

4.1.3.2. Group B: L1 Polish, L2 English, L3 German

Group B comprised 26 participants, native speakers of Polish who were students at the School of English, at Adam Mickiewicz University in Poznań, Poland, at the time of data collection. There were 20 female and 6 male participants and their mean age was 20.5 years (SD=1.1), ranging from 19 to 23 years old. For all of the participants English was their second language (L2) and German was their third language (L3), both in terms of chronology and the dominance of use. The level of proficiency in L2 English was fairly advanced (B2 - C1, according to CEFR), based on their practical English exam results, with an average length of training (L2_YFT) being 11.4 years (SD=2.7) and the age of onset (L2_AO) at 8.9 years old (SD=2.5). Their L2 English proficiency was evaluated on average as very good (M=5.1 on a 6 point scale, SD=0.4).

As for the L3 German, the participants' proficiency level ranged from elementary to lower intermediate (A2/B1 levels according to CEFR). The foreign language proficiency level was self-declared by the participants based on internal course placement assessment procedures. The average amount of time of formal training in German (L3_YFT) was 6.7 years (SD=3.7), whereas the mean age of onset of learning (L3_AOL) equalled 13 years (SD=4.1). The total number of foreign languages known by the participants totalled on average 2.8 (SD=0.8). Their self-evaluation of their general language competence in L3 German on a scale from 1-5 (1=very poor, 5=very good) equalled 2.4 (SD=0.9), similarly to the self-evaluation of their L3 pronunciation which was 2.3 (SD=0.8) corresponding to a category between satisfactory and good. As far as having stayed in German speaking countries is concerned, 42% of the participants had paid such a visit and the average length of stay for the whole population was 1.4 weeks (SD=3.4). The participants had undergone general linguistic training and L2 English pronunciation instruction, but no practical training of L3 German phonetic features was reported. The participants' profile for Group B is presented in Table 3.

Table 3. Participants' profile of Group B

N=26	L2 English				L3 German				
	Age	Prof	YFT	AOL	Prof	YFT	AOL	Eval max 5	Stay_L3 weeks
M	20.5	B2/C1	11.4	8.9	A2/B1	6.7	13	2.3	1.4
SD	1.1		2.7	2.5		3.7	4.1	0.8	3.4

YFT – years of formal training, AOL – age of onset of learning, Prof – proficiency (both as a proficiency level in CEFR and as assessed in a placement test (max 5), Eval – self-evaluation of L3 pronunciation, Stay – length of stay in L3 speaking countries (in weeks)

4.1.3.3. Group C: L1 Polish, L2 German, L3 English

Group C consisted of Polish students of the German philology at Adam Mickiewicz University in Poznań who had been learning English as their additional foreign language (L3). It comprised 34 participants (30 female and 4 male) with a mean age of 21.7 years (SD=3.6). They were very competent in L2 German at the B2/C1 proficiency levels, based on their practical German exam results, with 11.3 years of formal training (SD=2.8) and the mean age of onset being 9.9 (SD=2.8). Their L2 German proficiency was self-evaluated as very good (M=4.8 on a 5 point scale, SD=0.5).

English was the participants' third language (L3), which was acquired as a subsequent foreign language after L2 German. The proficiency level in L3 English ranged from A2 (elementary) to B1 (lower intermediate), based on internal course placement procedures, with an average length of exposure to English (L3_YFT) being 7.2 years (SD=2.8) and the mean age of onset of learning (L3_AOL) 14.0 years (SD=4.9). The participants' command of L3 English was self-evaluated on average as medium (M=2.4 on a 5 point scale, SD=0.8).

Only 23% of the participants had visited English speaking countries and the average length of stay for the whole population was 1.4 weeks (SD=3.6). Their self-evaluation of general English competence was at a moderate level, 2.4 out of 5 maximum (SD=0.8), and likewise they self-evaluated their L3 English pronunciation as medium (M=2.6, SD=0.8). The number of foreign languages known by the participants was on average 2.3 (SD=0.5). The participants had undergone general linguistic training during the course of their studies and pronunciation instruction in German, yet, no systematic practical training of L3 English phonetics was reported. A detailed participants' profile is presented in Table 4.

Table 4. Participants' profile of Group C

N=34	L2 German				L3 English				
	Age	Prof	YFT	AOL	Prof	YFT	AOL	Eval	Stay
M	21.7	B2/C1	11.3	9.9	A2/B1	7.2	14	2.6	1.4
SD	3.6		2.8	2.8		2.8	4.9	0.8	3.6

YFT – years of formal training, AOL – age of onset of learning, Prof – proficiency (both as a proficiency level in CEFR and as assessed in a placement test (max 5), Eval – self-evaluation of L3 pronunciation, Stay – length of stay in L3 speaking countries (in weeks).

4.1.3.4. Group D: L1 Polish, L2 French, L3 English

The participants in Group D involved Polish students of the French philology at Adam Mickiewicz University in Poznań who had been learning English as their additional foreign language (L3). Their total number was 29 (25 female and 4 male) and their mean age was 22.6 (SD=2.8). They were highly proficient in L2 French at the B2/C1 levels according to the classification of the Common European Framework of Reference (CEFR, 2001), with assessment based on their practical French exam results. The length of formal training in French (L2_YFT) was 6.8 (SD=4.2), whereas the mean age of onset of learning (L2_AO) equalled 15.9 years (SD=2.9). Their L2 French proficiency was self-evaluated as very good (M=4.4 on a 5 point scale, SD=0.9).

The participants have been learning English as their third language (L3), i.e. a subsequent foreign language after or parallel to French. Their proficiency level in L3 English ranged from B1 (lower intermediate) to B2 (intermediate), following internal course assessment procedures, with an average exposure to English (L3_YFT) being 11.1 years (SD=3.7). The mean age of onset of learning L3 English (L3_AOL) was 10.6 years (SD=3.5). Their L3 English proficiency was self-evaluated on average as rather good (M=3.8 on a 5 point scale, SD=0.8).

As far as having stayed in English speaking countries is concerned, 31% of the participants had paid such a visit and the average length of stay for the whole population was 1.9 weeks (SD=7.1). The participants were asked to evaluate their general competence in L3 English as well as their L3 pronunciation on a 5 point scale ranging from 1=very poor to 5=very good. Their self-evaluation of general English competence equalled 3.6 (SD=0.7). whereas they self-evaluated their L3 English pronunciation as medium (M=3.1, SD=0.7). The number of foreign languages known by the participants was on average 2.4 (SD=0.6). The participants had received general linguistic training during the course of their studies as well as pronunciation instruction in French; however, regular training of L3 English pronunciation was not provided. A detailed participants' profile is presented in Table 5.

Table 5. Participants' profile of Group D

N=29	L2 French				L3 English				
	Age	Prof	YFT	AOL	Prof	YFT	AOL	Eval	Stay
M	22.6	B2/C1	6.8	15.9	B1/B2	11.1	10.6	3.1	1.9
SD	2.8		4.2	2.9		3.7	3.5	0.7	7.1

YFT – years of formal training, AOL – age of onset of learning, Prof – proficiency (both as a proficiency level in CEFR and as assessed in a placement test (max 5), Eval – self-evaluation of L3 pronunciation, Stay – length of stay in L3 speaking countries (in weeks).

The background information collected by means of the language biographies resulted in a fairly homogenous participant profile for all the groups under investigation. As far as the L3 status of the respective languages is concerned, it was absolutely uncontroversial in Groups A, B and C, where the classification of a foreign language as a third language relied both on the sequence of acquisition as well as language proficiency and dominance (cf. Cenoz et al. 2001). However, in the case of Group D the status of English as a third language turned out to be more debatable. It was classified as L3 on the basis of the assessment of language dominance, with French being more dominant for the participants at the time of the data collection. On the other hand, in terms of language chronology, the subjects' first contact with a foreign language at school was with English rather than French, which is usually the case nowadays in the majority of Polish primary schools. Consequently, the descriptive statistics for the age of onset and length of formal training in Group D would favour English as the second language, yet due to the intensity of exposure to French and the recency of use, the language dominance had shifted and resulted in French taking priority as a dominant langue (thus L2) over English, which assumed the status of an L3.

After outlining the general assumptions of the three conducted investigations that constitute the empirical core of the present dissertation, I shall commence with Study 1 on accentedness ratings and proceed with a detailed description of the research design as well as a discussion of the findings. The remaining two studies will be outlined separately in the following chapters (5 and 6) and a global analysis entailing across study comparisons will be featured in Chapter 7.

4.2. Study 1: Research design

The first part of the series of parallel studies aimed at investigating cross-linguistic influence in L3 phonological acquisition based on perceptual ratings of foreign accentedness performed on samples of third language speech in various language combinations. It followed the basic design of previous pilot investigations into this field conducted by the present author (Wrembel 2012a, b) but with some further adaptations. The major objective was to further investigate the phenomenon of foreign accentedness as well as the sources and directionality of CLI in the acquisition of third language phonology.

The rationale behind selecting accentedness ratings was that although they have been widely used in studies on second language acquisition (e.g. Højen 2000, Piske et al. 2001, Gallardo del Puerto et al. 2007), they have not been applied extensively in the research on third language phonological acquisition. A notable exception was Hammarberg and Hammarberg's study (1993, 2005), in which L3 accent ratings were performed by three raters at two different stages of interlanguage development (see Section 3 for a detailed discussion). Specifically, the measure of perceived pronunciation performance was selected in order to generate foreign accent evaluation in a third language as well as to explore the L3 speakers' degree of comprehensibility and pronunciation accuracy judgements. Furthermore, it was expected to contribute to the investigation of the complexity of cross linguistic influence in TLA undertaken in the three different parts of the study.

The notion of a 'foreign accent' is used to refer to segmental and prosodic deviations from the native norms of pronunciation in a given language. In this study it is assessed holistically by raters on the basis of such rating parameters as the degree of foreign accent, speech intelligibility, and pronunciation correctness, i.e., the scales that are usually applied in foreign accent studies reported in the SLA literature (cf. Gallardo del Puerto et al. 2007; Piske et al. 2001). Cross-linguistic influence in L3 phonology is operationalised in the study as accentedness perceived by raters on the basis of their identification of the first language of the multilingual subjects based on the perceptual evaluation of speakers' L3 performance.

4.2.1. Aims and research questions

Study 1 aimed to investigate perceived pronunciation performance in the third language based on accentedness ratings. The ratings were designed to assess the accent strength in L3 speech samples, on the one hand, and the accent source, on the other. The former assessment involved three rating parameters, namely, the degree of foreign accent, comprehensibility and pronunciation correctness. The latter aimed to ascertain the perceived source of cross-linguistic influence in the third language based on L1 identification ratings. Further, the study was intended to explore potential relations between the perceived global accent ratings and extralinguistic variables, including the raters' and the speakers' characteristics.

To address the above-mentioned issues, the study posed the following research questions:

RQ 1: How are different rating parameters of the degree of foreign accent, comprehensibility and pronunciation correctness related in perceptual judgements performed on L3?
RQ 2: Do perceptual ratings in TLA exhibit similar patterns as those reported in SLA?
RQ 3: Do L1 identification patterns point to stronger native or non-native language influence on a perceived foreign accent in the L3?
RQ 4: Which factors, including raters' variables, have any bearing on the results?
RQ 5: Is there a correlation between the perceptual ratings and L3 proficiency level?

4.2.2. Participants and procedures

The study was conducted on a database of third language recordings collected by the author over the years 2012-13. The stimuli were recorded with the application of the Audition CS5.5 program as 16-bit mono files at a 41 000Hz sampling frequency. The subjects were recorded performing language tasks in the third language in two conditions: a read-on-your-own task and spontaneous speech; however, only the reading task was selected for perceptual ratings. Spontaneous speech performance proved challenging for lower level L3 speakers, and therefore data collection in the case of this task was not complete and it proved impossible to use this mode in designing the experiment.

For each language group (A, B, C, D) a pool of speech samples was selected randomly from the database of recordings of L3 English, French and German. The samples were verified for the quality of recording and shortened to 15-20 seconds excerpts. Additionally, several native speech samples in the languages under examination, recorded in identical conditions, were selected to serve as controls and added to the non-native pool. Next the prepared speech samples were randomised, coded and embedded into specially designed online questionnaires. Subsequently, they were subjected to foreign accent ratings performed by the raters of the respective languages. The ratings involved several parameters, including (1) the degree of the perceived foreign accent, (2) the evaluation of individual speakers' comprehensibility, (3) the evaluation of pronunciation correctness.

The experiment was designed on an online platform so that it could be performed by the raters via the Internet at their own convenience. The raters were recruited via a network of personal and professional contacts.

The online questionnaire was divided into two parts. The first one elicited the raters' personal information, including their native vs. non-native status, their proficiency level in the target language as well as their frequency of exposure to foreign-accented speech. The second part of the questionnaire consisted of identical subsections, each equipped with a recording and a set of 4 questions. The raters were asked to perform the following tasks:

1) rate the samples for an overall degree of foreign accent on a 7-point scale (1=strongly accented, 7=native-like accent);
2) evaluate the comprehensibility of the speaker (on a 7-point scale from 1=totally incomprehensible, 7=totally comprehensible)
3) assess the degree of pronunciation correctness (on a 7-point scale from 1=totally incorrect, 7=totally correct)
4) identify the speakers' L1 (i.e. select a language from an open list presented in the rating questionnaire).

Once the completed online questionnaire was submitted, the data was entered automatically into an excel sheet which facilitated the subsequent analysis of the results. Native control ratings were excluded from the actual analysis as they were identified correctly in nearly 100%.

There were two groups of participants involved in this study, which included:

(1) L3 learners who performed the L3 reading task and whose recordings were subject to the accent ratings, hereafter referred to as 'speakers',
(2) judges who evaluated perceptually the L3 samples by performing the online ratings, hereafter referred to as 'raters'.

In the case of the speakers, a language learning biography of each multilingual speaker was constructed in order to control better for various factors contributing to the complex nature of the cross linguistic influence. The relevant data were collected by means of a questionnaire administered individually after the recording session. The questions covered the following: (1) age of acquisition of each non-native language, (2) proficiency level in all non-native languages, (3) sequence of acquisition of all languages, (4) amount of formal instruction in the non-native languages (in years and hours per week), (5) natural exposure to the non-native languages (e.g. stays abroad), (6) number of languages known to the speaker, (7) metalinguistic awareness of positive or negative transfer from the non-native languages in L3 oral production. The analysis of the language biographies led to the construction of the participants' profiles, which are presented in an initial subsection for each of the four groups.

For the purpose of the present study, the original number of participants from each group had to be reduced and adapted to the requirements of the accent rating task. The main rationale behind this decision was to avoid fatigue effects on the part of the raters due to the time-consuming nature of the rating task. Therefore, the number of speech samples selected for the three rating experiments in L3 French, L3 German and L3 English was lower than the number of the respective participants in Groups A, B, C and D.

As far as the raters as concerned, they were asked to fill in a brief biodata questionnaire at the beginning of the online accentedness ratings that served as a control for the rater-specific variables including (1) native vs. non-native status, (2) proficiency level in the respective language, (3) phonetic training or lack thereof, (4) familiarity with accented speech, and

(5) knowledge of foreign languages. On the basis of the provided information, a separate profile was created for the raters in each of the studies A, B, C and D.

The author decided to include proficient non-native speakers of the respective languages as judges, apart from the native raters, although native speakers are generally believed to be naturally better at detecting a foreign accent (e.g., Piske et al. 2001). However, this assumption has recently been questioned by several scholars. Højen (2000), for instance, found evidence that non-native speakers are more sensitive to divergences from target language phonetic norms than natives, provided they have a distinct mental representation of the authentic pronunciation of the L2 sounds. Højen's claim is based on the perceptual magnet effect (cf. Kuhl & Iverson 1995), which states that natives are less sensitive to subcategorial phonetic differences between sounds close to a native phonetic prototype. Moreover, other foreign accent studies have provided evidence that non-native judges are as capable as native ones in detecting accentedness (cf. Flege 1988; Gallardo del Puerto et al. 2007, Wrembel 2010).

4.3. Results analysis Group A: L3 French

The results of perceptual pronunciation performance ratings in L3 will be presented separately for each group, followed by an across-groups comparison and a joint discussion.

The rating experiment on L3 French involved 30 speech samples including 28 samples of L3 French speech selected from the participants of Group A and 2 control samples of native French speech performing the same reading task. The original number of participants from Group A had to be slightly reduced and adapted to the requirements of the accent rating task.

Accent ratings were performed by a team of 22 judges that consisted of 10 native speakers of French (NS) and 12 Polish non-native speakers (NNS) with a near-native or advanced proficiency in French. Of all of the raters, 64% had previous phonetic training, whereas 36% did not. Their self-declared proficiency in French was fairly advanced (M=3.9 out of 5 maximum) and they had had on average rather frequent exposure to foreign accented French (M=4, on a 5 point scale, with 5 being the maximum).

4.3.1. Accentedness ratings – L3 French

The performed ratings involved several parameters, including (1) the degree of the perceived foreign accent, (2) the evaluation of individual speakers' comprehensibility, (3) the evaluation of pronunciation correctness. The mean value for each rating parameter is presented in Table 6. The ratings of native control samples were excluded from the analysis of results.

The first perceptual judgment task required the raters to assess the recorded L3 French samples for an overall degree of foreign accent on a 7-point scale (1=strongly accented, 7=native-like accent). The mean total rating was 3.1 (SD=1.7) and it turned out to be the lowest score of all the rating parameters. The comprehensibility ratings were significantly higher than those of foreign accentedness as the mean score was 4.9 (SD=1.8) on a 7-point scale (1=totally incomprehensible, 7=totally comprehensible). The pronunciation correctness rating had a mean score of 4.0 (1.7) on a 7-point scale (1=totally incorrect, 7=totally correct) and ranked as the medium judgements in between accentedness and comprehensibility scores.

Table 6. L3 French pronunciation performance ratings for Group A (on a 7-point scale; 1=lowest score, 7=highest score)

Ratings	N	Mean	SD	Median
Accent	616	3.1	1.7	3.0
Comprehensibility	616	4.9	1.8	5.0
Correctness	616	4.0	1.7	4.0

In order to compare the ratings as several independent samples the Kruskal-Wallis test was used as a nonparametric alternative to the one way ANOVA. The null hypothesis of the test is that all distribution functions are equal, while the alternative hypothesis is that at least one of the populations tends to yield larger values than at least one of the other populations. If the test is significant, multiple comparisons can be made between the samples, and the probability of each presumed non-difference is indicated.

As the performed Kruskal-Wallis test was found to be significant $H(2, n=84)=22.70888$, $p=0.0000$, multiple comparisons between the samples were performed. The pairwise comparisons showed statistically significant differences between the accentedness and comprehensibility ratings as well as between the accentedness and pronunciation correctness in L3 French (see Table 7). The highest ratings in the L3 French group were assigned by the judges for comprehensibility, whereas the lowest for accentedness as illustrated on the box-whisker plot (see Figure 1).

Table 7. Results of Kruskal-Wallis test and multiple comparisons tests for Group A

Kruskal-Wallis Test: H (2. n=84)=22.70888, p=0.0000			
p for multiple comparisons	Accent	Comprehensibility	Correctness
Accent	–	0.000006*	0.038929*
Comprehensibility	0.000006*	–	0.068004
Correctness	0.038929*	0.068004	–

*p<0.05

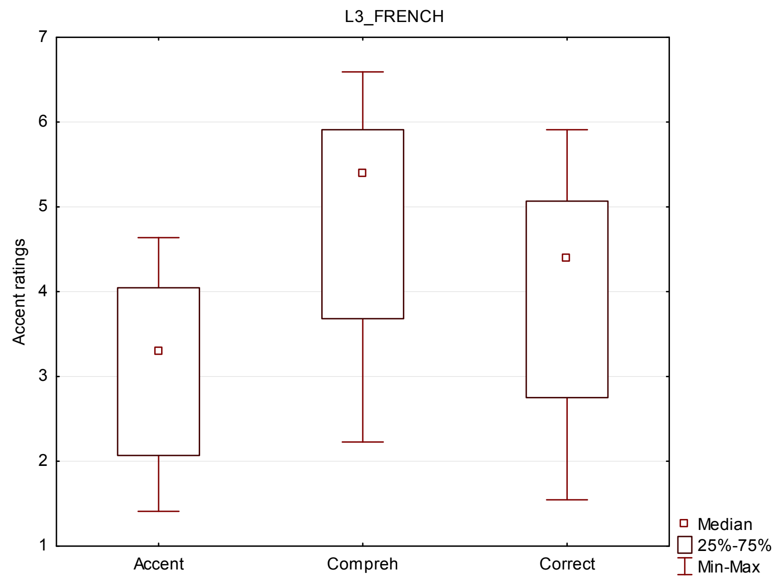


Figure 1. Box-whisker plot for accent ratings in L3 French

4.3.2. Intercorrelations between rating parameters in Group A

A Spearman rank correlation test was run to verify if the selected aspects of perceived pronunciation performance in L3 French are related to one another. The results of the test showed that there are strong positive correlations between comprehensibility and pronunciation correctness (R_S=0.78), between foreign accentedness and correctness (R_S=0.64) and a moderate correlation between foreign accentedness and comprehensibility (R_S=0.52), $p<0.05$ (see Table 8).

Table 8. Results of Spearman rank correlations between rating parameters for Group A

Variables	N	R_S	t(n-2)	p
Foreign accent & comprehensibility	616	**0.52**	15.26	0.00000*
Foreign accent & correctness	616	**0.64**	20.88	0.00000*
Comprehensibility & correctness	616	**0.78**	30.70	0.00000*

* $p<0.05$

4.3.3. L1 identification patterns in Group A

The source of cross-linguistic influence was operationalized in the study as perceived foreign accentedness related to the categorization of the speaker's first language, i.e. L1 identification, on the basis of L3 speech samples. In the L1 identification rating, the raters were to perform a forced-choice task by choosing one category from a provided list of languages including Polish, English, French, German, Russian, Spanish and others. The thus generated categorizations will be discussed as examples of correct L1 identifications, L1 identifications as the speakers' second language or incorrect L1 identifications as other languages. The percentages of respective categorizations will be presented separately for each group and interpreted as patterns of transfer sources, followed by an analysis of the raters' and speakers' variables and their impact on the performed ratings.

The percentage of correct L1 identifications as Polish totalled 36.2%, whereas the identification as L2 English was fairly low (9.4%) and other L1 identifications treated jointly reached the level of 54% including L1 identifications as German (10.4%), Spanish (10.2%), Russian (9.6%), French (0.5%), and other languages (23.7%) (see Figure 2). Although the correct L1 identification as Polish reached the highest score, it accounted for over one third of the responses, thus attesting an L1-accented speech in the L3 French performance. On the other hand, L2-accented speech was evidenced only in 9.4% of the cases; however, one may argue that this category could include also L1 identifications as German (10.4%), a language closely typologically related to English. Since the percentage of L1 identifications as languages other than the L1 or L2 treated as a joint category was rather high, it transpires that the raters were not able to correctly identify the participants' L1 on the basis of their L3 performance in approximately half of the cases. On the whole, the findings demonstrate complex patterns of interaction between the language systems of multilingual speakers and provide some further evidence against the claim that L1 is the only source of transfer in third or additional language acquisition.

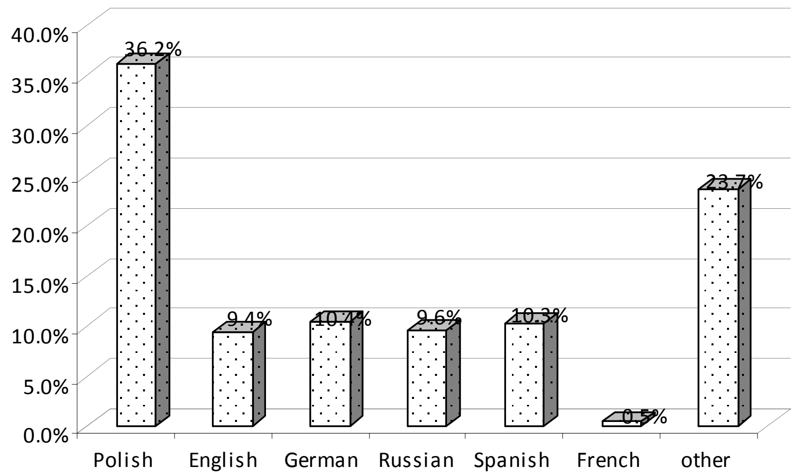


Figure 2. L1 identifications in Group A

4.3.4. Raters' variables – Group A

The generated ratings of perceived L3 pronunciation performance and L1 identification were subject to a series of analyses aimed at exploring the potential impact of the raters' variables, including (1) nativeness status, (2) language proficiency level, (3) and phonetic training. The results of the statistical testing will be presented in the subsequent sections.

The first analysis investigated an interdependency between the generated ratings and the raters' native vs. non-native status. A non-parametric Mann-Whitney test was applied to compare the differences in the assigned scores for various rating parameters (see Table 9). Only in the case of foreign accentedness did the scores differ significantly between native French and non-native raters (Z=2.53. p<0.05) implying that the native raters were more severe in their accent judgements (M=2.9, SD=1.5 for NS raters vs. M=3.3, SD=1.8 for NNS raters). The remaining differences did not prove significant, i.e. the scores assigned for comprehensibility judgements (M=4.9, SD=1.9 for NS raters vs. M=4.9, SD=1.8 for NNS raters) and pronunciation correctness (M=4.0, SD=1.7 for NS raters vs. M=3.9, SD=1.7 for NNS raters) were comparable irrespective of the raters' native or non-native status.

Table 9. Rating scores for native vs. non-native raters in Group A

Raters' status	L3 French											
	Accent				Comprehensibility				Correctness			
	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian
NNS	336	3.3*	1.8	3.0	336	4.9	1.8	5.0	336	3.9	1.7	4.0
NS	280	2.9*	1.5	3.0	280	4.9	1.9	5.0	280	4.0	1.7	4.0
Total	616	3.1	1.7	3.0	616	4.9	1.8	5.0	616	4.0	1.7	4.0

* p<0.05

The second analysis investigated if there is any relation between the rating scores and the raters' proficiency level in French. To this end, a Spearman rank correlation was applied. Weak negative correlations were observed for the raters' proficiency level in French and accentedness scores (R_S=-0.13) as well as comprehensibility scores (R_S=-0.12), p<0.05 (see Table 10). It follows that raters with higher proficiency levels assigned lower scores for accentedness and comprehensibility, but this correlation was not observed in the case of the pronunciation correctness scores.

Table 10. Spearman rank correlation for raters' proficiency level and ratings scores in Group A

Pairs of variables	N	R_S	t(n-2)	p
French proficiency & Accent	616	-0.13	-3.33	0.00091*
French proficiency & Comprehensibility	616	-0.12	-2.94	0.00336*
French proficiency & Correctness	616	0.00	-0.04	0.96808

* p<0.05

In the third analysis, a non-parametric Mann-Whitney test was performed to assess if there was any influence of the variable of phonetic training on the ratings. No significant differences between the groups of raters who did or did not have any previous phonetic training were observed for accentedness (M=3.1, SD=1.7 for phonetically trained raters vs. M=3.0, SD=1.6 for non-trained ones) and comprehensibility ratings (M=4.8, SD=1.9 for phonetically trained raters vs. M=5.0, SD=1.7 for non-trained ones). However, the difference between the groups proved significant in the case of pronunciation correctness ratings (Z=-2.6, $p<0.05$), with phonetically trained raters scoring on average M=3.8, (SD=1.7); and non-trained raters M=4.2, (SD=1.5) (see Table 11). The results imply that those raters who were not phonetically trained tended to assign higher ratings for L3 pronunciation correctness.

Table 11. Rating scores for phonetically trained vs. untrained raters in Group A

Phonetic training	L3 French											
	Accent				Comprehensibility				Correctness			
	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian
YES	392	3.1	1.7	3.0	392	4.8	1.9	5.0	392	3.8*	1.7	4.0
NO	224	3.0	1.6	3.0	224	5.0	1.7	5.0	224	4.2*	1.5	4.0
Total	616	3.1	1.7	3.0	616	4.9	1.8	5.0	616	4.0	1.7	4.0

* $p<0.05$

Further analyses concerned the potential impact of raters' variables on the results of the L1 identification task. As far as the interaction between L1 identification and raters' nativeness status, a Chi-square test demonstrated a statistically significant correlation between them (χ^2=33.04. df=6. $p<0.01$). A detailed analysis showed that the percentage of correct identification of speech samples as L1 Polish was higher for non-native raters who were Polish themselves (44.9%) as compared to the native French raters (25.7%). In the case of native raters, the identification as German (15.4%) or other languages (27.9%) generated higher percentages than in the other group of raters (6.3% and 20.2% respectively). Interestingly, the identification as English (i.e. the participants' L2) was nearly identical for both groups of raters (9.5% vs. 9.3%). Figure 3 illustrates the L1 identification patterns separately for native vs. non-native raters.

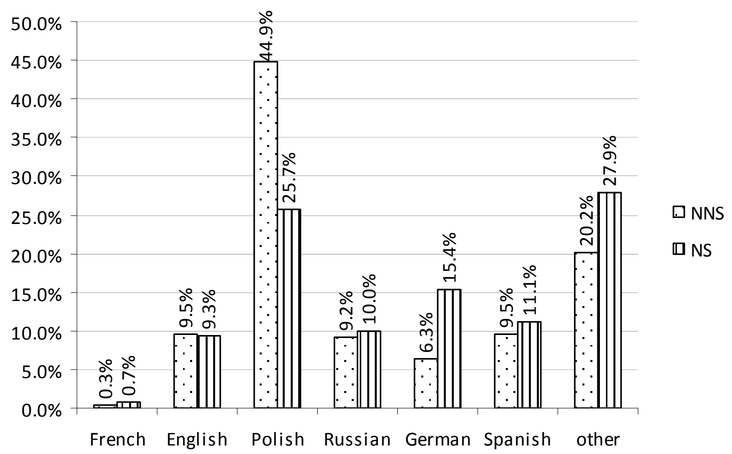


Figure 3. L1 identification vs. raters' native status in Group A

Finally, the investigation into the relation between L1 identification and the raters' phonetic training proved to be statistically significant (χ^2=53.3, df=6, p<0.05). Phonetically trained raters had a higher percentage of correct L1 identifications as Polish (45.4%) than untrained raters (20%), whereas the latter outscored the former in the L1 identification as German (17.4% vs. 6.4%) and other languages (27.2% vs. 21.7%). The findings demonstrate that phonetically trained raters are better at identifying correctly the speakers' L1, whereas the raters without phonetic training tend to be more inclined towards incorrect L1 identifications as other foreign languages (see Figure 4).

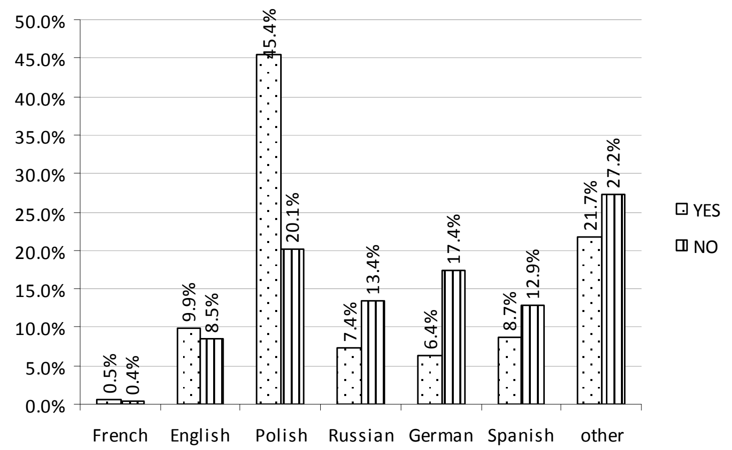


Figure 4. L1 identification vs. raters' phonetic training – Group A

4.3.5. Interraters' reliability and consistency

The interraters' reliability was assessed by means of the Cronbach alpha coefficient of internal consistency between raters. It was demonstrated to be very high for all the rating parameters (accentedness α=0.95, comprehensibility α=0.95, correctness α=0.93, L1 identification α=0.88).

The consistency of the obtained ratings was measured by means of Kendall's coefficient of concordance. Kendall's coefficient of concordance was found to be moderate to high for the three rating parameters: accentedness (W=0.54), comprehensibility (W=0.68), and correctness (W=0.70); however, the ratio was low for the L1 identification (W=0.12) (see Table 12).

Table 12. Interraters' reliability and consistency in Group A

L3 French	Cronbach alpha	Kendall's coefficient of concordance
Accent	0.95	0.54
Comprehensibility	0.95	0.68
Correctness	0.93	0.70
L1 Identification	0.88	0.12

4.3.6. Correlation analysis for accentedness ratings and speakers' variables

Spearman rank correlation tests were performed to investigate correlations between perceived pronunciation performance (rated as accentedness, comprehensibility and correction) and the speakers' variables. Significant strong to moderate correlations ($p<0.01$) were found for several variables including the length of formal training in L3 French, the onset age of learning of L3 French, the participants' age, self-evaluation of one's L3 pronunciation and general competence, the L2 English and L3 French proficiency levels as well as the year of studies (see Table 13).

The length of formal training in L3 French (L3_YFT) correlated strongly with all the pronunciation rating parameters, verifying at the same time the internal validity of the performed accent, comprehensibility and pronunciation correctness ratings. The following correlations were observed to hold for formal training in L3 French: L3_YFT & Accent (R_S=0.72); L3_YFT & Comprehensibility (R_S=0.69); L3_YFT & Correctness (R_S=0.68).

Negative strong correlations were found between the age of onset of L3 French (L3_AOL) and all the measures of perceived L3 performance, indicating that the earlier the participants started learning L3 French, the better they were rated on accentedness, comprehensibility and pronunciation correctness. The specific results of rank correlations were as follows: L3_AOL & Accent (R_S=-0.71); L3_AOL & Comprehensibility (R_S=-0.66); L3_AOL & Correctness (R_S=-0.66).

The age of the speakers correlated moderately with the ratings, indicating that the older they were, the better they were rated on accent, comprehensibility and pronunciation correctness. The observed correlations included: Age & Accent (R_S=0.46); Age & Comprehensibility (R_S=0.57); Age & Correctness (R_S=0.47).

The speakers' self-evaluation of L3 pronunciation as well as general competence in L3 were found to correlate strongly to moderately with all the measures of perceived L3 French performance. The self-evaluation of one's L3 competence correlated with Accent (R_S=0.68); with Comprehensibility (R_S=0.57) and with Correctness (R_S=0.59). Similarly, self-evaluation of one's L3 pronunciation correlated with Accent (R_S=0.67); with Comprehensibility (R_S=0.64) and with Correctness (R_S=0.64).

Very strong correlations were found between the speakers' proficiency level in L3 French (L3_Prof) and all the rating parameters, indicating that the more proficient they were in L3 French, the better scores they received on accentedness, comprehensibility and correctness ratings, which further confirmed the internal validity of the performed ratings. The specific results of the correlations for L3 proficiency were as follows: with Accent (R_S=0.83); with Comprehensibility (R_S=0.84); with Correctness (R_S=0.82). Moreover, the L2 English proficiency level (L2_Prof) correlated moderately with all L3 pronunciation performance ratings, generating the following correlations: L2_Prof & Accent (R_S=0.5); L2_Prof & Comprehensibility (R_S=0.54); L2_Prof & Correctness (R_S=0.52).

Finally, the year of studies correlated moderately with all the perceptual measures of the L3 French performance, i.e. Year & Accent (R_S=0.53); Year & Comprehensibility (R_S=0.56); Year & Correctness (R_S=0.53).

Table 13. Spearman rank correlations between L3 accentedness ratings and speakers' variables for Group A L3 French

Pairs of variables for L3 French	N	R_S	t(n-2)	p
L2_YFT & Accent	28	-0.06	-0.29	0.776589
L2_YFT & Comprehensibility	28	0.03	0.16	0.874929
L2_YFT & Correctness	28	0.01	0.04	0.971076
L2_AOL & Accent	28	0.31	1.67	0.106306
L2_AOL & Comprehensibility	28	0.33	1.78	0.086952
L2_AOL & Correctness	28	0.30	1.60	0.121433
L3_YFT & Accent	28	**0.72**	5.25	**0.000017**
L3_YFT & Comprehensibility	28	**0.69**	4.92	**0.000041**
L3_YFT & Correctness	28	**0.68**	4.73	**0.000069**
L3_AOL & Accent	28	**-0.71**	-5.13	**0.000024**
L3_AOL & Comprehensibility	28	**-0.66**	-4.45	**0.000144**
L3_AOL & Correctness	28	**-0.66**	-4.43	**0.000152**
N_TOTAL_Ln & Accent	28	-0.20	-1.05	0.304611
N_TOTAL_Ln & Comprehensibility	28	-0.32	-1.73	0.094691
N_TOTAL_Ln & Correctness	28	-0.32	-1.73	0.095692
Stay in L3 & Accent	11	0.06	0.17	0.872113
Stay in L3 & Comprehensibility	11	-0.04	-0.12	0.906625
Stay in L3 & Correctness	11	0.08	0.24	0.814716
Age & Accent	28	**0.46**	2.65	**0.013555**
Age & Comprehensibility	28	**0.57**	3.54	**0.001517**
Age & Correctness	28	**0.47**	2.73	**0.011215**
Eval L3 comp & Accent	28	**0.68**	4.72	**0.000070**
Eval L3 comp & Comprehensibility	28	**0.57**	3.58	**0.001376**
Eval L3 comp & Correctness	28	**0.59**	3.71	**0.000997**
Eval L3 pron & Accent	28	**0.67**	4.66	**0.000083**
Eval L3 pron & Comprehensibility	28	**0.64**	4.29	**0.000221**
Eval L3 pron & Correctness	28	**0.64**	4.30	**0.000211**
L2_Prof & Accent	28	**0.50**	2.93	**0.006985**
L2_Prof & Comprehensibility	28	**0.54**	3.30	**0.002804**
L2_Prof & Correctness	28	**0.52**	3.07	**0.004921**

L3_Prof & Accent	28	**0.83**	7.62	**0.000000**
L3_Prof & Comprehensibility	28	**0.84**	7.94	**0.000000**
L3_Prof & Correctness	28	**0.82**	7.32	**0.000000**
Year & Accent	28	**0.53**	3.17	**0.003880**
Year & Comprehensibility	28	**0.56**	3.45	**0.001916**
Year & Correctness	28	**0.53**	3.15	**0.004111**

YFT – years of formal training, AOL – age of onset of learning, Prof – proficiency, N_total_Ln – number of foreign languages known, Eval L3 pron – self-evaluation of L3 pronunciation, Eval L3 com – self-evaluation of general competence in L3, Stay – length of stay in L3 speaking countries, Year – year of studies.

4.4. Results analysis – Group B: L3 German

For the purpose of the present rating study, 25 speech samples of L3 German were selected from the participants of Group B as well as 3 control samples of native Germans performing the same reading task.

Accent ratings were performed by a team of 20 judges that consisted of 16 native speakers of German (NS) and 4 Polish non-native speakers (NNS) with a near-native or advanced proficiency in German. Of all of the raters, 55% had previous phonetic training, whereas the remaining 45% did not. Their self-declared proficiency in English was very advanced (M=4.7 out of 5 maximum, SD=0.8) and they had on average frequent exposure to foreign accented German speech (M=4.0, out of 5 maximum, SD=1.0). The raters declared their knowledge of, on average, three foreign languages including English, Polish, French, Spanish, Russian and others.

4.4.1. Accentedness ratings – L3 German

The performed perceptual judgment tasks involved three parameters and required the raters to assess the recorded L3 German samples on a 7-point scale for an overall degree of foreign accentedness (1=strongly accented, 7=native-like accent), comprehensibility (1=totally incomprehensible, 7=totally comprehensible) and pronunciation correctness (1=totally incorrect, 7=totally correct).

In Group B the mean score for the rating of foreign accentednes was 3.1 (SD=1.7) and it was the lowest score of all the rating parameters. In turn, the comprehensibility rating proved significantly higher as the mean score was 4.9 (SD=1.5). Finally, pronunciation correctness was ranked as intermediate between the other two, with a mean score of 4.0 (SD=1.4), (see Table 14).

Table 14. L3 German accentedness ratings for Group B

Ratings	N	Mean	SD	Median
Accent	500	3.1	1.7	3.0
Comprehensibility	500	4.9	1.5	5.0
Correctness	500	4.0	1.4	4.0

A Kruskal-Wallis test was used as a method for comparing several independent samples and it proved to be significant H(2, n=75)=35.17529, p=0.0000. Consequently, multiple pairwise comparisons between the scores were performed pointing to statistically significant differences between all the rating parameters involved (see Table 15). The highest ratings in Group B were assigned by the judges for comprehensibility, whereas the lowest was for accentedness in L3 German. For a box-whisker plot presentation see Figure 5.

Table 15. Results of Kruskal-Wallis test and multiple comparisons tests for Group B

Kruskal-Wallis Test: H(2, n=75)=35.17529, **p=0.0000**			
p for multiple comparisons	Accent	Comprehensibility	Correctness
Accent	–	0.000000*	0.017986*
Comprehensibility	0.000000*	–	0.004475*
Correctness	0.017986*	0.004475*	–

*p<0.05

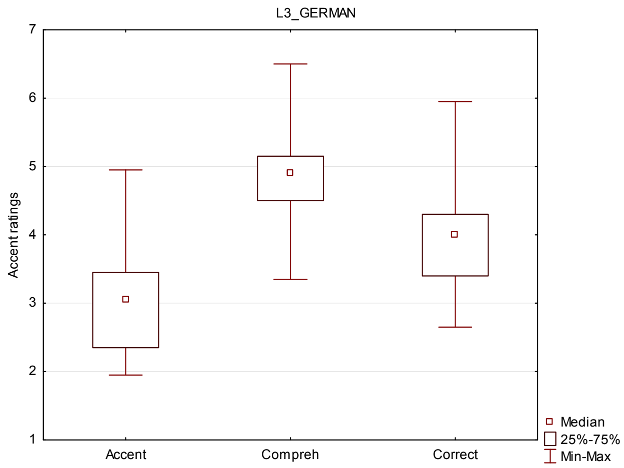


Figure 5. Box-whisker plot for pronunciation ratings in L3 German

4.4.2. Intercorrelations between rating parameters in Group B

In the subsequent analysis, Spearman rank correlation tests were run to verify if the selected aspects of perceived pronunciation performance in L3 German are interrelated. The results demonstrated that there are statistically significant strong positive correlations between foreign accentedness and pronunciation correctness (R_S=0.79) as well as comprehensibility and pronunciation correctness (R_S=0.71), whereas the correlation between foreign accentedness and comprehensibility proved moderate (R_S=0.57), $p<0.05$ (see Table 16). It follows that the more a native like accent was assigned to the samples, the more they were evaluated as comprehensible and correct in terms of pronunciation.

Table 16. Spearman rank correlations between rating parameters in L3 German

Variables	N	R_S	t(n-2)	p
Foreign accent & comprehensibility	500	**0.57**	15.64	0.00000*
Foreign accent & correctness	500	**0.79**	28.56	0.00000*
Comprehensibility & correctness	500	**0.71**	22.43	0.00000*

* $p<0.05$

4.4.3. L1 identification patterns in Group B – L3 German

On the whole, the percentage of correct L1 identification as Polish, reached the level of 41.6%, with identification as L2 English being 10.6%, and other L1 identifications treated jointly 47.8% including: identification as Russian (11.8%), Spanish (6.6%), French (6.6%) and other languages (22.4%) (see Figure 6). In the case of Group B, the level of correct identifications as L1 Polish is high, thus clearly pointing to some evidence of an L1-accented speech in the L3 German performance. The second hypothesized outcome would favour L2 English as a potential source of cross-linguistic influence; however, the evidence for L2-accented performance in L3 German is rather weak with only a 10.6% level, which is comparable to that of other foreign languages identification scores.

All in all, the results indicate rather complex patterns of interaction between the language systems of a multilingual speaker and provide some evidence against the claim that L1 is the only source of transfer in third or additional language acquisition. Moreover, identifications as other languages were rather high, which may suggest that the raters were not able to correctly identify the participants' L1 on the basis of their L3 performance in nearly half of the cases.

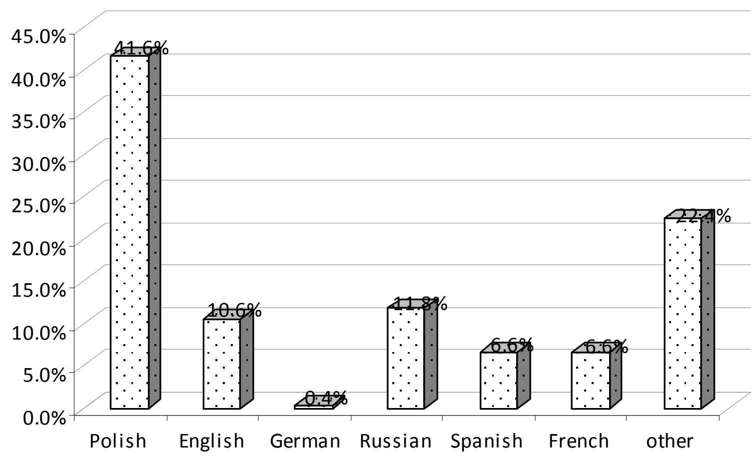


Figure 6. L1 identifications for L3 German in Group B

4.4.4. Raters' variables – Group B

The first investigated raters' variable was that of native vs. non-native status. Non-parametric Mann-Whitney tests were performed to compare the L3 German ratings with respect to the raters' nativeness status. The difference between the two groups proved to be statistically significant only in the case of comprehensibility (M=5.0 for NS vs. M=4.5 for NNS; Z=-2.5, p<0.05), indicating that native speaker raters were less severe in their ratings and found the accented samples more comprehensible. For the remaining parameters of foreign accentedness and pronunciation correctness, no significant differences in the rating scores were found between native and non-native raters (see Table 17).

Table 17. Rating scores for native vs. non-native raters in Group B

Raters' status	L3 German											
	Accent				Comprehensibility				Correctness			
	N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median
NNS	100	3.1	1.8	3.0	100	4.5*	1.6	5.0	100	3.7	1.6	4.0
NS	400	3.1	1.6	3.0	400	5.0*	1.5	5.0	400	4.0	1.4	4.0
Total	500	3.1	1.7	3.0	500	4.9	1.5	5.0	500	4.0	1.4	4.0

* p<0.05

To account for any potential impact of the raters' proficiency level in German on their respective ratings of L3 pronunciation performance, another series of Spearman rank correlations was performed. A significant, yet weak positive correlation was observed between the raters' proficiency level and comprehensibility scores (R_S=0.11, p<0.05). It follows that raters with higher proficiency levels in German assigned higher scores for comprehensibility for the L3 German samples. No other interdependencies were found between this raters' variable and the remaining rating parameters (see Table 18).

Table 18. Spearman rank correlations for Group B

Pair of variables	N	R_S	t(n-2)	p
German proficiency & Accent	500	0.02	0.52	0.60527
German proficiency & Comprehensibility	500	0.11	2.45	0.01446*
German proficiency & Correctness	500	0.09	1.96	0.05015

* p<0.05

In order to assess any potential influence of another raters' variable, i.e. phonetic training on the performed ratings, a non-parametric Mann-Whitney test was calculated. No significant differences between the groups of raters who did or did not have any previous phonetic training were observed for foreign accentedness (M=3.0, SD=1.6 for phonetically trained raters vs. M=3.3, SD=17 for untrained ones). Nonetheless, the between the group difference proved statistically significant (Z=-2.2, p<0.05) in the comprehensibility ratings, i.e. phonetically untrained raters (M=5.0, SD=1.5) found speakers to be more comprehensible than the trained ones (M=4.7, SD=1.5). Significantly different results were also observed for the pronunciation correctness ratings (Z=-2.4, p<0.05) in which phonetically trained raters assigned more severe ratings (M=3.8, SD=1.3) than raters without any phonetic training (M=4.1, SD=1.5), (see Table 19).

Table 19. Rating scores for phonetically trained vs. untrained raters in Group B

Phonetic training	L3 German											
	Accent				Comprehensibility				Correctness			
	N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median
YES	275	3.0	1.6	3.0	275	4.7*	1.5	5.0	275	3.8*	1.3	4.0
NO	225	3.3	1.7	3.0	225	5.0*	1.5	5.0	225	4.1*	1.5	4.0
Total	500	3.1	1.7	3.0	500	4.9	1.5	5.0	500	4.0	1.4	4.0

* p<0.05

Subsequent analyses focused on the interdependencies between the raters' variables and the results of the L1 identification task. A conducted Chi-square test showed a statistically significant relation between L1 identification and the raters' native vs. non-native status (χ^2=40.57. df=6. p<0.01). As was demonstrated in a detailed analysis, the percentage of correct identification of speech samples as L1 Polish was higher for native German raters (46.3%) as compared to non-native raters who were themselves Polish (23%). On the other hand, non-native raters tended to identify the speech samples as English, which was actually the speakers' L2, more frequently (22%) than the native raters (7.8%). Detailed L1 identification patterns for native vs. non-native German raters are illustrated in Figure 7.

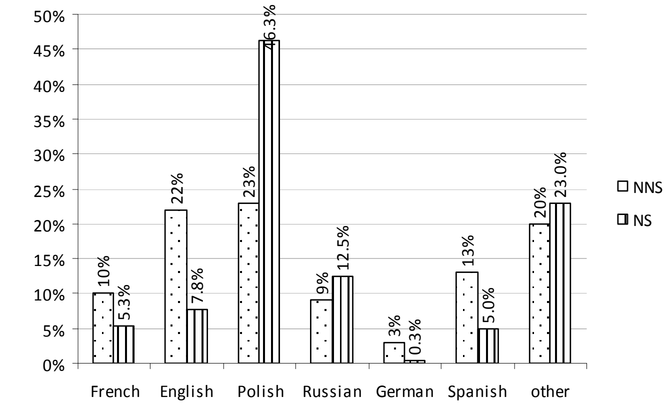


Figure 7. L1 identification vs. raters' native status in Group B

The investigation into the relation between L1 identification patterns and raters' phonetic training demonstrated to be statistically significant (χ^2=12.9, df=6, p<0.05). Phonetically trained raters had a higher percentage of L1 identifications as English (13%) or other languages (25%) than untrained raters (7.6% and 19% respectively), whereas phonetically untrained raters outscored the former in correct L1 identification as Polish (44% vs. 39.6%). The results point to a slight tendency for phonetically untrained raters to identify better speaker's L1, whereas the raters with phonetic training tend to be more inclined towards identifications as English, which is the speakers' L2 or as other foreign languages (see Figure 8).

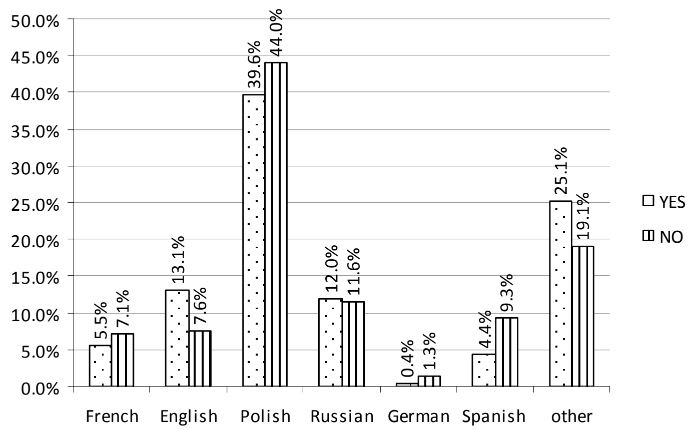


Figure 8. L1 identification vs. raters' phonetic training for Group B

4.4.5. Interraters' reliability and consistency

The Cronbach alpha coefficient of internal consistency between raters was used to evaluate the interraters' reliability. The scales reliability proved excellent (>0.9) for accentedness, comprehensibility and correctness and very high (0.7-0.9) for L1 identification. The coefficients for specific rating parameters were as follows: accentedness α=0.98, comprehensibility α=0.94, correctness α=0.95, L1 identification α=0.86.

The consistency of the obtained ratings was measured by means of Kendall's coefficient of concordance. Kendall's coefficient of concordance was found to be moderate for the three rating parameters: accentedness (W=0.48), comprehensibility (W=0.39), and correctness (W=0.45); however, the ratio was very low for L1 identification (W=0.06) (see Table 20).

Table 20. Interraters' reliability and consistency

L3 German	Cronbach alpha	Kendall's coefficient of concordance
Accent	0.978	0.482
Comprehensibility	0.944	0.386
Correctness	0.952	0.452
L1 Identification	0.864	0.060

4.4.6. Correlation analysis for accentedness ratings and speakers' variables

In order to investigate potential correlations between perceived pronunciation performance rated as accentedness, comprehensibility and correction, and the participants' variables, further Spearman rank correlations were performed.

Significant moderate to strong correlations ($p<0.05$) were found for several variables, including years of formal training in L3 German, age of onset of L3 learning, L3 proficiency and self-evaluation of one's L3 pronunciation as well as general competence (see Table 21).

The length of formal training in L3 German (L3_YFT) correlated moderately with all the rating parameters, verifying at the same time the internal validity of the performed accent, comprehensibility and pronunciation correctness ratings. The specific results of the rank correlations were as follows: L3_YFT & Accent (R_S=0.61); L3_YFT & Comprehensibility (R_S=0.47); L3_YFT & Correctness (R_S=0.58).

There were negative moderate correlations between the age of onset of L3 German (L3_AOL) and all the measures of perceived L3 performance, indicating that the earlier the participants started learning L3 German, the better they were rated on accentedness, comprehensibility and pronunciation correctness. The following correlations were observed to hold for the age of onset in L3: L3_AOL & Accent (R_S=-0.52); L3_AOL & Comprehensibility (R_S=-0.42); L3_AOL & Correctness (R_S=-0.49). Only one correlation was found for the age of onset of L2 English (L2_AOL) and it was a moderate correlation with pronunciation correctness in L3, i.e. L2_AOL & Correctness (R_S=0.42).

The speakers' self-evaluation of their L3 pronunciation as well as general competence correlated moderately with all the measures of perceived L3 pronunciation performance. In turn, self-evaluation of one's L3 competence correlated with Accent (R_S=0.46); with Comprehensibility (R_S=0.41); and with Correctness (R_S=0.46). Similar correlation patterns were observed for the self-evaluation of their own L3 pronunciation, including Accent (R_S=0.61); Comprehensibility (R_S=0.4); and Correctness (R_S=0.58).

Finally, strong correlations were found between the speakers' proficiency level in L3 German and all three rating parameters, including L3_Proficiency & Accent (R_S=0.76); L3_Proficiency & Comprehensibility (R_S=0.65); L3_Proficiency & Correctness (R_S=0.78).

Table 21. Spearman rank correlations between accentedness ratings and speakers' variables for Group B

Pair of variables L3_GERMAN	N	R_S	t(n-2)	p
L2_YFT & Accent	25	-0.35	-1.82	0.082268
L2_YFT & Comprehensibility	25	-0.21	-1.01	0.321874
L2_YFT & Correctness	25	-0.35	-1.77	0.089845
L2_AOL & Accent	25	0.39	2.01	0.056567
L2_AOL & Comprehensibility	25	0.33	1.68	0.105878
L2_AOL & Correctness	25	**0.42**	2.20	**0.038039**
L3_YFT & Accent	25	**0.61**	3.64	**0.001355**
L3_YFT & Comprehensibility	25	**0.47**	2.56	**0.017444**

L3_YFT & Correctness	25	**0.58**	3.39	**0.002511**
L3_AOL & Accent	25	**-0.52**	-2.90	**0.007992**
L3_AOL & Comprehensibility	25	**-0.42**	-2.25	**0.034657**
L3_AOL & Correctness	25	**-0.49**	-2.66	**0.013885**
N_TOTAL_Ln & Accent	25	0.07	0.33	0.747979
N_TOTAL_Ln & Comprehensibility	25	0.10	0.50	0.625281
N_TOTAL_Ln & Correctness	25	0.08	0.36	0.721364
Stay in L3 & Accent	11	0.09	0.27	0.790087
Stay in L3 & Comprehensibility	11	-0.06	-0.20	0.849458
Stay in L3 & Correctness	11	0.10	0.30	0.773369
Age & Accent	25	-0.17	-0.82	0.421301
Age & Comprehensibility	25	-0.16	-0.79	0.435388
Age & Correctness	25	-0.20	-1.00	0.327544
Eval L3 comp & Accent	25	**0.46**	2.51	**0.019727**
Eval L3 comp & Comprehensibility	25	**0.41**	2.13	**0.043649**
Eval L3 comp & Correctness	25	**0.46**	2.47	**0.021231**
Eval L3 pron & Accent	25	**0.61**	3.67	**0.001286**
Eval L3 pron & Comprehensibility	25	**0.40**	2.11	**0.045643**
Eval L3 pron & Correctness	25	**0.58**	3.42	**0.002366**
L2_Prof & Accent	25	0.10	0.49	0.627517
L2_Prof & Comprehensibility	25	0.00	0.01	0.993718
L2_Prof & Correctness	25	0.03	0.13	0.896712
L3_Prof & Accent	25	**0.76**	5.55	**0.000012**
L3_Prof & Comprehensibility	25	**0.65**	4.11	**0.000428**
L3_Prof & Correctness	25	**0.78**	6.03	**0.000004**
Year & Accent	25	-0.02	-0.10	0.918717
Year & Comprehensibility	25	-0.07	-0.35	0.730344
Year & Correctness	25	-0.11	-0.55	0.586554

4.5. Results analysis – Groups C and D: L3 English

Accentedness ratings on L3 English were prepared jointly for Groups C and D since both groups had English as their L3 and the decision was motivated by a greater ecological validity of the study.

The rating involved 36 speech samples (15-20 second long) including 17 samples from Group C, 16 samples from Group D, and 3 control samples featuring an English, French and German native speakers speaking English.

The accent ratings were performed by a team of 23 judges that consisted of 12 native speakers of English (NS), and 11 Polish non-native speakers (NNS) with a near-native or advanced proficiency in English. Of all of the raters, 87% had previous phonetic training, whereas the remaining 13% did not. Their self-declared proficiency in English was very advanced (M=4.4 out of 5 maximum, SD=0.7) and they had had on average very frequent exposure to foreign accented English speech (M=4.7, out of 5 maximum, SD=0.6). The raters declared the knowledge of, on average, three foreign languages including Polish, French, Spanish, German, Russian and others.

4.5.1. Accentedness ratings – L3 English

The performed ratings involved several parameters, including (1) the degree of the perceived foreign accent, (2) the evaluation of individual speakers' comprehensibility, (3) the evaluation of pronunciation correctness. The mean value for each rating parameter is presented in Table 22. Native control samples were excluded from the analysis of results. The subsequent analysis will present the results jointly for L3 English samples, followed by a comparison of subcomponent groups C and D.

In the first perceptual judgement task the raters assessed the recorded L3 English samples for an overall degree of foreign accent on a 7-point scale (1=strongly accented, 7=native-like accent). This assessment turned out to present the lowest score of all the rating parameters; the mean total rating for L3 English being 3.2 (SD=1.4). The foreign accentedness rating was lower in Group C with L2 German (M=2.9, SD=1.3) than in Group D with L2 French, where its level was considerably higher (M=3.5, SD=1.4).

As expected, the comprehensibility ratings were significantly higher than those of the foreign accentedness as the total mean score for L3 English was 4.8 (SD=1.6) on a 7-point scale (1=totally incomprehensible, 7=totally comprehensible). As in the previous rating parameter, the score for Group D was also better in terms of comprehensibility (M=5.1, SD=1.5) than for Group C (M=4.5, SD=1.6).

The third parameter, namely the pronunciation correctness rating, was ranked intermediate between the other two. The total mean correctness ratings for L3 English was 3.5 (SD=1.3) on a 7-point scale (1=totally incorrect, 7=totally correct), with Group D (M=3.8, SD=1.3) outscoring Group C (M=3.3, SD=1.3).

All in all, the general rating pattern was preserved across the groups, with foreign accentedness being the most severely judged parameter followed by the correctness ratings. Comprehensibility, on the other hand, was consistently rated much higher, indicating that foreign accented speech does not necessarily lead to incomprehensibility; in spite of the fact that speech samples may be perceived as accented, they are still fairly comprehensible.

Table 22. Accentedness ratings for L3 English (Groups C and D)

Group	Parameter	N	Mean	SD	Median
L3 English total	Accent	759	3.2	1.4	3.0
	Comprehensibility		4.8	1.6	5.0
	Correctness		3.5	1.3	3.0
Group C with L2 German	Accent	391	2.9	1.3	3.0
	Comprehensibility		4.5	1.6	5.0
	Correctness		3.3	1.3	3.0
Group D with L2 French	Accent	368	3.5	1.4	3.0
	Comprehensibility		5.1	1.5	5.0
	Correctness		3.8	1.3	4.0

(1 – lowest score, 7 – highest score)

A subsequent analysis involved a comparison of the results of two subgroups of L3 English accentedness ratings, i.e. Group C with L2 German, and Group D with L2 French. The performed Manna-Whitney's tests pointed to statistically significant differences between the groups for all the rating parameters. As far as foreign accentedness is concerned, the L3 English_L2 German group was judged to be more accented (M=2.9, SD=1.3) than the L3 English_L2 French one (M=3.5, SD=1.4) and the difference was statistically significant ($Z=6.1$, $p<0.01$). Likewise, comprehensibility ratings demonstrated the same trend of a statistically significant difference ($Z=4.7$, $p<0.01$), with the L3 English_L2 French group (M=5.1, SD=1.5) rated as more comprehensible than the L3 English_L2 German (M=4.5, SD=1.6).

Finally, a similar pattern was displayed in the pronunciation correctness scores (Z=5.6, p<0.01) with the L3 English_L2 German participants judged on average as significantly less correct (M=3.3, SD=1.3) than the participants of the L3 English_L2 French group (M=2.9, SD=1.3) (see Table 23).

Table 23. Accentedness ratings for L3 English – subgroup comparison

Ratings	Group D			Group C			Manna-Whitney test	
	N	Mean	SD	N	Mean	SD	Z	p
Accentedness	368	3.5	1.4	391	2.9	1.3	6.42	0.0000*
Comprehensibility	368	5.1	1.5	391	4.5	1.6	4.69	0.0000*
Correctness	368	3.8	1.3	391	3.3	1.3	5.59	0.0000*

In order to compare the accent ratings as several independent parameters, non-parametric Kruskal-Wallis tests were performed for subgroups C and D of the L3 English accentedness rating.

In Group C the results of the test were found to be significant H(2, n=51)=29.97336, p=0.0000, thus multiple comparisons between the samples were made. The pairwise comparisons showed statistically significant differences between accentedness and comprehensibility ratings as well as between comprehensibility and pronunciation correctness (see Table 24). Following the trends in all the other language repertoires, the highest scores in this group were assigned by the raters for comprehensibility, whereas the lowest were for accentedness (see Figure 9 for a box-whisker plot illustration).

Table 24. Results of Kruskal-Wallis test and multiple comparisons tests for Group C

Kruskal-Wallis Test: H (2, n=51)=29.97336, **p=0.0000**			
p for multiple comparisons	Accent	Comprehensibility	Correctness
Accent	–	0.000000*	0.405571
Comprehensibility	0.000000*	–	0.000412*
Correctness	0.405571	0.000412*	–

*p<0.05

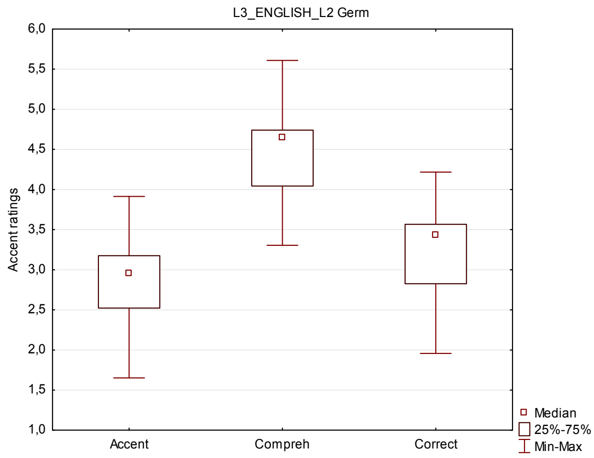


Figure 9. Box-whisker plot for accent ratings in Group C

As far as subgroup D of the accentedness ratings on L3 English is concerned, the performed Kruskal-Wallis test was also found to be significant H(2, n=48)=28.65299, p=0.0000, therefore multiple comparisons between the samples were conducted. Identically to the previous group, statistically significant differences were found between accentedness and comprehensibility, as well as between comprehensibility and pronunciation correctness ratings for Group D (see Table 25). As illustrated in Figure 10, the raters evaluated L3 comprehensibility the highest, whereas foreign accentedness was rated the lowest in Group D, thus following the trends exhibited in all the other investigated groups.

Table 25. Results of Kruskal-Wallis test and multiple comparisons tests for Group D

Kruskal-Wallis Test: H (2, n=48)=28.65299 p=0.0000			
p for multiple comparisons	Accent	Comprehensibility	Correctness
Accent	–	0.000001*	0.647775
Comprehensibility	0.000001*	–	0.000302*
Correctness	0.647775	0.000302*	–

*p<0.05

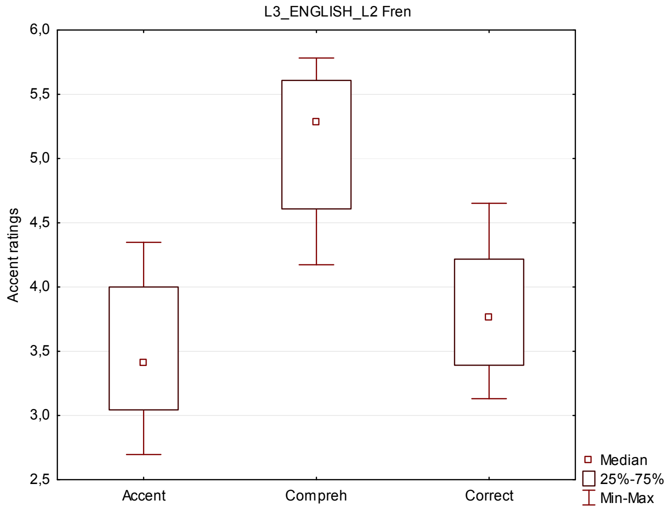


Figure 10. Box-whisker plot for accent ratings in Group D

4.5.2. Intercorrelations between rating parameters in L3 English (Groups C and D)

In order to assess whether the selected aspects of perceived pronunciation performance in L3 English are interrelated, a Spearman rank correlation test was performed. Strong positive correlations were found in all the groups between foreign accentedness and pronunciation correctness in L3 English ratings treated jointly (R_S=0.81), whereas the correlations were moderate between comprehensibility and pronunciation correctness (R_S=0.56) and foreign accentedness and comprehensibility (R_S=0.49), $p<0.05$. A detailed analysis of subcomponent groups C and D demonstrated nearly identical patterns of correlations as the ones reported in the joint analysis of the L3 English ratings (see Table 26).

Table 26. Results of Spearman rank correlations for L3 English (Groups C and D)

Group	Pairs of variables	N	R_S	t(n-2)	p
L3 English_total	Foreign accent & comprehensibility	759	0.49	15.53	0.00000*
	Foreign accent & correctness	759	0.81	37.55	0.00000*
	Comprehensibility & correctness	759	0.56	18.54	0.00000*

Group D with L2 French	Foreign accent & comprehensibility	368	0.47	10.13	0.00000*
	Foreign accent & correctness	368	0.78	23.94	0.00000*
	Comprehensibility & correctness	368	0.51	11.43	0.00000*
Group C with L2 German	Foreign accent & comprehensibility	391	0.48	10.78	0.00000*
	Foreign accent & correctness	391	0.82	27.99	0.00000*
	Comprehensibility & correctness	391	0.57	13.53	0.00000*

*$p<0.05$

4.5.3. L1 identification patterns in L3 English

The analysis of L1 identification patterns in L3 English samples was conducted separately for the two subcomponents groups. In Group C the highest L1 identification scores were attributed correctly to Polish (29.2%), followed by identification as other foreign languages (18.7%) and as German (17.7%), which is actually the speakers' L2. Similar patterns were observed in Group D, with the correct identification as Polish ranking as first of all the ratings (37.8%), followed by identification as other languages (21.5%) and as French (14.1%), which was the speakers' L2 in this group (see Figure 11).

On the basis of a performed Chi-square test, the relation between L1 identification and the two subgroups was found to be statistically significant ($\chi^2=33.34$. $df=6$. $p<0.01$). A detailed analysis demonstrated that the percentage of correct identification as L1 Polish was higher for Group D (37.8% vs. 29.2%), whereas the identification as German was considerably higher for Group C with L2 German (17.6% vs. 5.7%). The difference in L1 identification as French also approached significance with Group C with L2 French outscoring Group D (14.3% vs. 10.2%). Interestingly, the identification as French or German was found to be related to the group L2 profile, i.e. the identification scores were higher for the respective L2 groups. This implies that the participants' L2 exerted some influence on the perceived foreign accentedness of their L3 English.

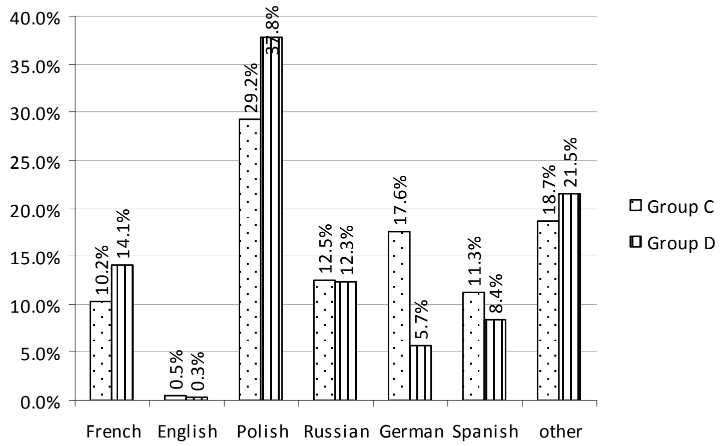


Figure 11. L1 identification for Groups C and D

4.5.4. Raters' variables – Groups C and D

Raters' variables including native status, phonetic training and language proficiency level were subject to a subsequent analysis in order to verify their impact on the performed L3 English ratings. The differences in the rating scores for various parameters were calculated by means of non-parametric Mann-Whitney tests.

As far as native vs. non-native status is concerned, the raters' scores differed significantly only in the case of the foreign accentedness ratings in L3 English_total (Z=2.2. $p<0.05$) with the native raters being slightly more severe in their accent judgements (M=3.1, SD=1.3) than non-native raters (M=3.3, SD=1.4). No such difference was attested in the separate analysis of subcomponent groups C and D. The remaining parameters in L3 English total ratings did not demonstrate any significant differences and they proved to be comparable irrespective of the raters' native or non-native status, i.e. comprehensibility judgements (M=4.9, SD=1.7 for NS raters vs. M=4.8, SD=1.4 for NNS raters), and pronunciation correctness (M=3.5, SD=1.3 for NS raters vs. M=3.5, SD=1.4 for NNS raters). Similar patterns were exhibited in a separate analysis of Groups C and D (see Table 27).

Table 27. L3 English rating scores for native vs. non-native raters for Groups C and D

Status	Group analysis											
	Group D				Group C				L3 English_total			
	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian
Foreign accent												
NNS	176	3.6	1.4	3.0	187	3.0	1.4	3.0	363	3.3*	1.4	3.0
NS	192	3.4	1.3	3.0	204	2.7	1.3	3.0	396	3.1*	1.3	3.0
Comprehensibility												
NNS	176	5.0	1.4	5.0	187	4.5	1.4	5.0	363	4.8	1.4	5.0
NS	192	5.1	1.6	5.0	204	4.6	1.7	5.0	396	4.9	1.7	5.0
Correctness												
NNS	176	3.8	1.4	4.0	187	3.3	1.4	3.0	363	3.5	1.4	4.0
NS	192	3.8	1.2	4.0	204	3.2	1.2	3.0	396	3.5	1.3	3.0

*p<0.05

In order to explore any relation between the rating scores and the raters' proficiency level in English, Spearman rank correlation tests were calculated. No statistically significant correlations were found between this rater variable and their assigned scores for accentedness, comprehensibility and pronunciation correctness in any of the groups, i.e. L3 English ratings treated jointly as well as separate analyses for Groups C and D.

Non-parametric Mann-Whitney tests were performed with the view to verifying any potential influence of the variable of phonetic training on the raters' perceptual judgements of L3 English. No significant differences between phonetically trained and untrained raters were found for two of the rating parameters, namely, accentedness and pronunciation correctness in all the groups under investigation (i.e. L3 English total, subcomponent Groups C and D).

Nevertheless, the difference proved significant in the case of the comprehensibility ratings for the joint analysis of L3 English ($Z=3.5$, $p<0.05$) and that of Group D ($Z=3.2$, $p<0.05$), with phonetically trained raters assigning higher scores for comprehensibility ($M=4.9$ and $M=5.2$ for the respective groups) than non-trained raters ($M=4.3$ and $M=4.5$ respectively) (see Table 28).

Table 28. L3 English rating scores for phonetically trained vs. untrained raters for Groups C and D

Phonetic training	Group analysis											
	Group D				Group C				L3 English_total			
	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian	N	Mean	SD	Me-dian
Foreign accent												
YES	320	3.5	1.3	3.0	340	2.8	1.3	3.0	660	3.2	1.3	3.0
NO	48	3.2	1.9	3.0	51	2.9	1.8	2.0	99	3.0	1.8	3.0
Comprehensibility												
YES	320	5.2*	1.5	5.0	340	4.6	1.6	5.0	660	4.9*	1.6	5.0
NO	48	4.5*	1.2	5.0	51	4.2	1.4	4.0	99	4.3*	1.3	5.0
Correctness												
YES	320	3.8	1.3	4.0	340	3.2	1.3	3.0	660	3.5	1.3	3.0
NO	48	4.0	1.5	4.5	51	3.5	1.5	3.0	99	3.7	1.5	4.0

*p<0.05

The L1 identification rates for L3 English analysed jointly are illustrated separately for native vs. non-native raters in Figure 12.

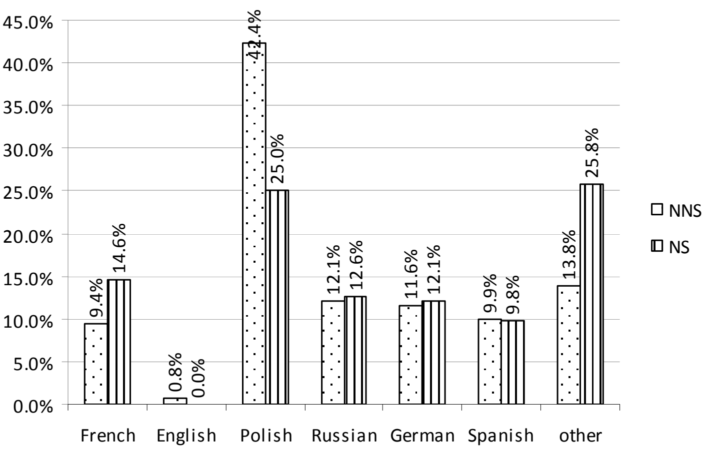


Figure 12. L1 identification for native vs. non-native raters of L3 English (C and D treated jointly)

A Chi-square test demonstrated a statistically significant relation between L1 identification and the raters' nativeness status (χ^2=40.2, p<0.01). As transpires from a more detailed analysis, the percentage of correct identification of speech samples as L1 Polish was higher for non-native raters who were Polish speakers themselves (42.4%) as compared to native English raters (25%). On the other hand, the identification as other languages generated considerably higher percentage scores in the group of native raters (25.8%) than the non-native ones (13.8%). The remaining identification categories remained fairly comparable irrespective of the raters' native speaker status.

A separate statistical analysis for Group D did not point to any significant relation between L1 identification and the raters' native language status. Still, there was a visible tendency for a higher percentage of L1 identifications as Polish in non-native raters (45.5% vs. 30.7%) and for identification as other languages (24.5% vs. 18.2%) in native raters.

As far as Group C is concerned, statistically significant correlations were found between the raters' native language status and L1 identification (χ^2=39.3, p<0.05). Parallel tendencies were observed as in the other group (D) with the correct L1 identification as Polish being higher for non-native raters who were Polish themselves (39.6% vs. 19.6%), and L1 identification as other languages being higher for the native raters (27% vs. 9.6%) (see Figure 13).

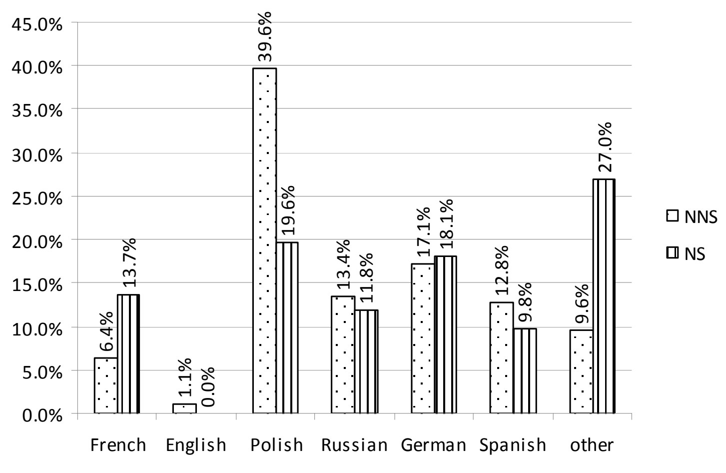


Figure 13. L1 identification for native vs. non-native raters for Group C

The investigation into the relation between L1 identification and raters' phonetic training proved to be significant for the L3 English scores treated jointly (χ^2=15.9, df=6, p<0.05). Phonetically trained raters generated a higher percentage of correct L1 identifications as Polish than untrained raters (35.2% vs. 21.2%), whereas the latter outscored the former in identification as German (19.2% vs. 10.8%) or, to a lesser extent, French (16.2% vs. 11.5%) (see Figure 14). These findings indicate a tendency for phonetically trained raters to identify more correctly the speaker's L1, whereas the raters without phonetic training are more inclined towards identifications as German or French, which are the speakers' respective second languages (L2).

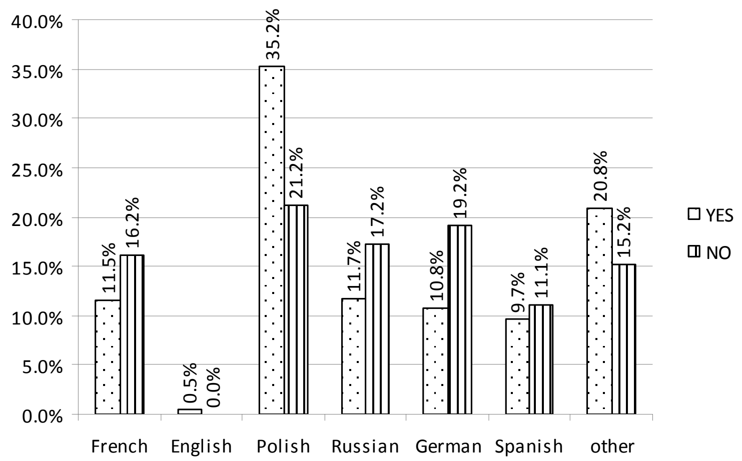


Figure 14. L1 identification vs. raters' phonetic training in L3 English

4.5.5. Interraters' reliability and consistency

The interraters' reliability was assessed by means of the Cronbach alpha coefficient of internal consistency between raters. It was found to be very high for all the rating parameters for L3 English treated jointly (accentedness α=0.97, comprehensibility α=0.98, correctness α=0.97, L1 identification α=0.87) as well as for Groups C and D analysed separately. Detailed results for all the groups are presented in Table 29. On the whole, the scales reliability proved excellent (>0.9) for accentedness, comprehensibility and correctness, and high (0.7-0.9) for L1 identification.

Table 29. Cronbach alpha coefficients for L3 English

Cronbach alpha	Group D	Group C	L3 English_ total
Accent	0.953	0.962	0.977
Comprehensibility	0.966	0.959	0.980
Correctness	0.943	0.963	0.976
L1 Identification	0.767	0.749	0.869

The consistency of the obtained ratings was calculated by means of Kendall's coefficient of concordance. Kendall's coefficient of concordance was found to be moderate to weak for the three rating parameters (W=0.4-0.3), but the ratio was low for L1 identification (W<0.2). For instance, in the joint analysis of L3 English ratings, the coefficient of concordance was moderate for accentedness (W=0.38), comprehensibility (W=0.38) and correctness (W=0.38) and weak for L1 identification (W=0.12). Table 30 illustrates detailed results for the groups under investigation.

Table 30. Kendall's concordance coefficient

Kendall's concordance coefficient	Group D	Group C	L3 English_ total
Accent	0.295	0.376	0.379
Comprehensibility	0.329	0.351	0.376
Correctness	0.281	0.411	0.378
L1 Identification	0.150	0.084	0.117

p<0.05

4.5.6. Correlation analysis for accentedness ratings and speakers' variables

In order to investigate potential correlations between perceived pronunciation performance rated as accentedness, comprehensibility and correctness and the participants' variables, further Spearman rank correlations were calculated for the two subgroups separately.

In Group C statistically significant correlations (p<0.05) were found only in the case of one variable, namely, the age of onset of L3 learning (L3_AOL). There were negative moderate correlations between the age of onset of L3 English and two measures of perceived L3 performance, indicating that the earlier the participants started learning L3 English, the better they were rated on accentedness and comprehensibility. The following correlations were observed to hold for the age of onset of L3 English: L3_AOL & Accent (R_S=-0.55); L3_AOL & Comprehensibility (R_S=-0.52). The remaining speakers' variables did not correlate significantly with any of the ratings parameters in Group C (see Table 31).

In turn, in Group D statistically significant correlations ($p<0.05$) were found also only in the case of one variable, namely, the year of studies, which correlated strongly to moderately with all the rating parameters, including accentedness, comprehensibility and pronunciation correction in L3. This means that the participants with more years of studies scored better in all the ratings. The specific results of rank correlations were as follows: Year & Accent (R_S=-0.56); Year & Comprehensibility R_S=-0.60); Year & Correctness (R_S=-0.67). No further correlations were found to hold between the speakers' variables and perceived pronunciation performance measures in L3 English in Group D (see Table 32).

Table 31. Spearman rank correlations between accentedness ratings and speakers' variables for Group C

Pairs of variables for L3 English_L2 German	N	R_S	t(n-2)	p
L2_YFT & Accent	17	0.13	0.51	0.617100
L2_YFT & Comprehensibility	17	0.12	0.48	0.636686
L2_YFT & Correctness	17	0.24	0.94	0.363398
L2_AOL & Accent	17	-0.13	-0.50	0.623236
L2_AOL & Comprehensibility	17	-0.07	-0.27	0.792654
L2_AOL & Correctness	17	-0.05	-0.21	0.835460
L3_YFT & Accent	17	0.44	1.87	0.080556
L3_YFT & Comprehensibility	17	0.48	2.09	0.053674
L3_YFT & Correctness	17	0.46	1.99	0.065473
L3_AOL & Accent	17	**-0.55**	-2.58	**0.021025**
L3_AOL & Comprehensibility	17	**-0.52**	-2.36	**0.032554**
L3_AOL & Correctness	17	-0.41	-1.73	0.103431
N_TOTAL_Ln & Accent	17	0.23	0.90	0.381810
N_TOTAL_Ln & Comprehensibility	17	0.37	1.54	0.144816

N_TOTAL_Ln & Correctness	17	0.28	1.14	0.270618
Stay in L3 & Accent	2	–	–	–
Stay in L3 & Comprehensibility	2	–	–	–
Stay in L3 & Correctness	2	–	–	–
Age & Accent	17	0.03	0.13	0.899294
Age & Comprehensibility	17	0.10	0.41	0.690384
Age & Correctness	17	0.22	0.88	0.390740
Eval L3 comp & Accent	17	0.39	1.66	0.117082
Eval L3 comp & Comprehensibility	17	0.37	1.57	0.138233
Eval L3 comp & Correctness	17	0.29	1.18	0.257584
Eval L3 pron & Accent	17	0.34	1.38	0.186800
Eval L3 pron & Comprehensibility	17	0.30	1.21	0.243498
Eval L3 pron & Correctness	17	0.15	0.60	0.558634
L2_Prof & Accent	17	0.17	0.66	0.520397
L2_Prof & Comprehensibility	17	0.20	0.79	0.441804
L2_Prof & Correctness	17	0.29	1.16	0.262300
L3_Prof & Accent	17	0.40	1.67	0.115759
L3_Prof & Comprehensibility	17	0.36	1.50	0.154665
L3_Prof & Correctness	17	0.37	1.53	0.147394
Year & Accent	17	0.09	0.35	0.733947
Year & Comprehensibility	17	0.18	0.73	0.478617
Year & Correctness	17	0.24	0.95	0.356487

Table 32. Spearman rank correlations between accentedness ratings and speakers' variables for Group D

Pairs of variables L3 English_L2 French	N	R_S	t(n-2)	p
L2_YFT & Accent	16	0.29	1.15	0.267903
L2_YFT & Comprehensibility	16	0.48	2.04	0.060588
L2_YFT & Correctness	16	0.42	1.74	0.103965
L2_AOL & Accent	16	-0.19	-0.72	0.483750
L2_AOL & Comprehensibility	16	-0.41	-1.71	0.110198
L2_AOL & Correctness	16	-0.30	-1.17	0.259868
L3_YFT & Accent	16	0.44	1.84	0.086969

L3_YFT & Comprehensibility	16	0.49	2.10	0.054450
L3_YFT & Correctness	16	0.48	2.07	0.057239
L3_AOL & Accent	16	-0.12	-0.47	0.644846
L3_AOL & Comprehensibility	16	-0.13	-0.47	0.642121
L3_AOL & Correctness	16	-0.05	-0.20	0.847968
N_TOTAL_Ln & Accent	16	-0.44	-1.83	0.088838
N_TOTAL_Ln & Comprehensibility	16	-0.35	-1.41	0.181141
N_TOTAL_Ln & Correctness	16	-0.45	-1.91	0.077031
Stay in L3 & Accent	4	0.26	0.38	0.741801
Stay in L3 & Comprehensibility	4	0.77	1.73	0.225403
Stay in L3 & Correctness	4	0.77	1.73	0.225403
Age & Accent	16	-0.04	-0.14	0.888756
Age & Comprehensibility	16	0.06	0.22	0.826799
Age & Correctness	16	0.06	0.22	0.826925
Eval L3 comp & Accent	16	0.31	1.20	0.249984
Eval L3 comp & Comprehensibility	16	0.30	1.19	0.254084
Eval L3 comp & Correctness	16	0.24	0.94	0.364242
Eval L3 pron & Accent	16	0.10	0.39	0.705444
Eval L3 pron & Comprehensibility	16	0.20	0.77	0.456665
Eval L3 pron & Correctness	16	0.17	0.66	0.517752
L2_Prof & Accent	16	0.13	0.50	0.625615
L2_Prof & Comprehensibility	16	0.39	1.60	0.131984
L2_Prof & Correctness	16	0.34	1.33	0.203274
L3_Prof & Accent	16	0.07	0.25	0.804847
L3_Prof & Comprehensibility	16	0.36	1.43	0.174569
L3_Prof & Correctness	16	0.22	0.85	0.407936
Year & Accent	16	**-0.56**	-2.50	**0.025326**
Year & Comprehensibility	16	**-0.60**	-2.81	**0.014039**
Year & Correctness	16	**-0.67**	-3.38	**0.004507**

4.6. Joint analysis of accentedness ratings for all the groups

A global analysis was performed for all the four groups treated jointly (A, B, C, D).

4.6.1. L3 Accent ratings

4.6.1.1. Foreign accentedness

The mean scores for the overall degree of foreign accentedness in L3 on a 7-point scale (1=strongly accented, 7=native-like accent) was M=3.1 (SD=0.8) and it was identical for Groups A and B (M=3.1). The scores were found to be slightly lower in case of L3 Group C (M=2.9) and higher for Group D (M=3.5) (see Table 33).

Table 33. Descriptive statistics for foreign accent ratings

Group	Accent			
	N	Mean	SD	Median
A: L3_FRENCH	28	3.1	1.0	3.3
B: L3_GERMAN	25	3.1	0.8	3.0
C: L3_ENGLISH_L2 German	17	2.9	0.6	3.0
D: L3_ENGLISH_L2 French	16	3.5	0.5	3.4
Total	86	3.1	0.8	3.1

The performed Kruskal-Wallis test demonstrated that the probability level was not significant (H (3, n=86)=5.904730 p=0.12), indicating that the rating scores for foreign accentedness in L3 did not differ significantly across the four investigated language groups (see Figure 15).

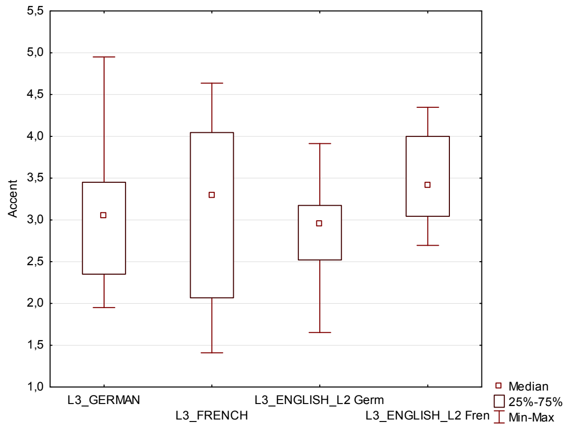


Figure 15. Box-whisker plot for accent ratings in all language groups

4.6.1.2. Comprehensibility ratings

The comprehensibility ratings in L3 (1=totally incomprehensible, 7=totally comprehensible) proved significantly higher than accentedness as the mean score was 4.9 (SD=0.9), and it was identical for Groups A and B (M=4.9), although the median was higher for Group A (5.4). The average scores were slightly lower in the case of Group C (M=4.5) and higher for Group D (M=5.1) (see Table 34).

Table 34. Descriptive statistics for comprehensibility ratings

Group	Comprehensibility			
	N	Mean	SD	Median
A: L3_FRENCH	28	4.9	1.3	5.4
B: L3_GERMAN	25	4.9	0.8	4.9
C: L3_ENGLISH_L2 German	17	4.5	0.7	4.7
D: L3_ENGLISH_L2 French	16	5.1	0.6	5.3
Total	86	4.9	0.9	4.9

As the probability level of the computed Kruskal-Wallis test for comprehension was not found to be significant (H (3, n=86)=4.6 p=0.2), the differences between comprehensibility ratings in the four investigated language groups were not significant (see Figure 16).

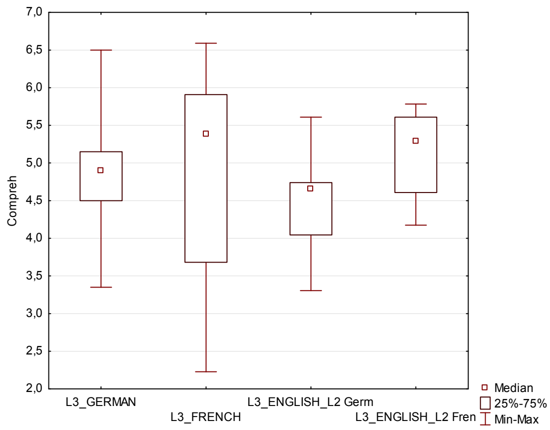


Figure 16. Box-whisker plot for comprehensibility ratings in all language groups

4.6.1.3. Pronunciation correctness

Pronunciation correctness in L3 was ranked as intermediate between the other two ratings with a mean score of 3.8 (SD=0.9), where 1=totally incorrect and 7=totally correct. As in the previous ratings, the mean scores were identical for Groups A and B (M=4.0), although the median was higher for Group A (4.4). The correctness scores were lower in the case of Group C (M=3.3) and slightly lower for Group D (M=3.8) (see Table 35).

Table 35. Descriptive statistics for pronunciation correctness ratings

Group	Correctness			
	N	Mean	SD	Median
B: L3_GERMAN	25	4.0	0.8	4.0
A: L3_FRENCH	28	4.0	1.3	4.4
C: L3_ENGLISH_L2 German	17	3.3	0.6	3.4
D: L3_ENGLISH_L2 French	16	3.8	0.5	3.8
Total	86	3.8	0.9	3.7

The Kruskal-Wallis test was used for comparing several independent samples (as a nonparametric alternative to the one way ANOVA). As the performed test was found to be significant H (3, N=86)=8.41 p=0.04, multiple comparisons between the samples were made. The pairwise comparisons demonstrated a statistically significant difference between Groups A and C. The lowest ratings for pronunciation correctness were assigned by the judges to Group C, whereas the highest were assigned to Group A (see Table 36, Figure 17).

Table 36. Results of Kruskal-Wallis test and multiple comparisons tests for L3 pronunciation correctness - total * $p<0.05$

Kruskal-Wallis Test: H (3, N=86)=8.408500 p=0.0383				
p for multiple comparisons	Group B	Group A	Group C	Group D
Group B	–	1.000000	0.112306	1.000000
Group A	1.000000	–	**0.033100***	1.000000
Group C	0.112306	**0.033100***	–	0.419192
Group D	1.000000	1.000000	0.419192	–

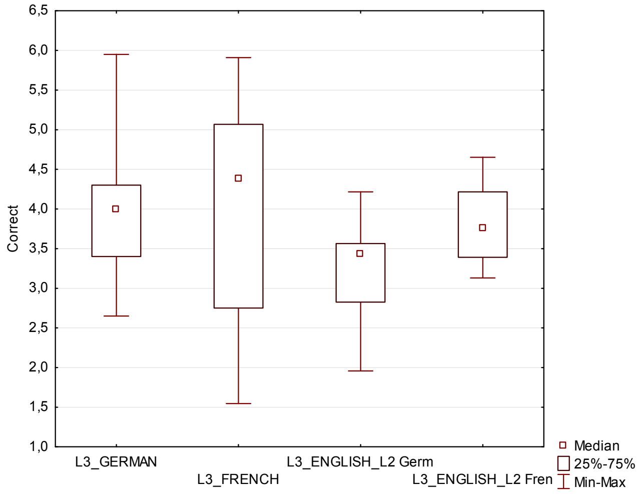


Figure 17. Box-whisker plot for pronunciation correctness ratings in all language groups

4.6.2. Correlations between accentedness ratings and speakers' variables – joint analysis

Spearman rank correlations were performed to investigate correlations between the perceived pronunciation performance (rated as accentedness, comprehensibility and correction) and the participants' variables for all the groups treated jointly. This joint analysis has a considerable statistical power as it was performed on a sample of 86 L3 participants. Significant strong to moderate correlations ($p<0.01$) were found for several variables, including the length of formal training in L3, the onset age of L3 learning, proficiency levels in L2 and L3 as well as self-evaluation of the participants' own L3 pronunciation and general L3 competence (see Table 37).

The joint correlational analysis pointed to the L3 proficiency level as the variable exhibiting the strongest overall correlations with all the parameters of perceived pronunciation performance in L3. It follows that the more proficient the participants in their respective L3, the better scores they received on accentedness, comprehensibility and correctness ratings, which further confirmed the internal validity of the performed ratings. The correlations for L3 proficiency were as follows: L3_Proficiency & Accent ($R_S=0.66$); L3_Proficiency & Comprehensibility ($R_S=0.60$); L3_Proficiency & Correctness ($R_S=0.57$).

Further, self-evaluation of one's L3 pronunciation as well as general L3 competence were also found to correlate strongly to moderately with all the measures of perceived L3 performance. Self-evaluation of one's L3 competence correlated with Accent (R_S=0.56), with Comprehensibility (R_S=0.47), and with Correctness (R_S=0.43). Similar patterns of interdependencies were exhibited by self-evaluation of L3 pronunciation which correlated with Accent (R_S=0.54); with Comprehensibility (R_S=0.48) and with Correctness (R_S=0.43). It follows that self-evaluation was a strong predictor of pronunciation performance.

The length of formal training in L3 (L3_YFT) correlated moderately with all the rating parameters, verifying at the same time the internal validity of the performed accent, comprehensibility and pronunciation correctness ratings. The following correlations were observed to hold for the years of formal L3 training: L3_YFT & Accent (R_S=0.53); L3_YFT & Comprehensibility (R_S=0.4); L3_YFT & Correctness (R_S=0.36).

Negative moderate correlations were found between the age of onset of L3 learning (L3_AOL) and all the measures of perceived L3 performance, indicating that the earlier the participants started learning their respective L3, the better they were rated on accentedness, comprehensibility and pronunciation correctness. The specific results of the rank correlations for the age of L3 onset included: L3_AOL & Accent (R_S=-0.49); L3_AOL & Comprehensibility (R_S=-0.36); L3_AOL & Correctness (R_S=-0.34).

There was a weak correlation between the participants' age and foreign accentedness ratings in a joint group analysis, indicating that older participants tended to receive better accentedness scores (Age & Accent R_S=0.22).

Moreover, the L2 proficiency level correlated weakly with two rating parameters, namely, with Comprehensibility (R_S=0.23) and with Correctness (R_S=0.26).

Table 37. Spearman rank correlations between accentedness ratings and speakers' variables – joint analysis

Pairs of variables_ JOINT ANALYSIS	N	R_S	t(n-2)	p
L2_YFT & Accent	86	-0.12	-1.11	0.269196
L2_YFT & Comprehensibility	86	0.02	0.16	0.874218
L2_YFT & Correctness	86	0.08	0.70	0.485674

L2_AOL & Accent	86	**0.27**	2.59	**0.011267**
L2_AOL & Comprehensibility	86	0.15	1.41	0.162461
L2_AOL & Correctness	86	0.06	0.53	0.598928
L3_YFT & Accent	86	**0.53**	5.71	**0.000000**
L3_YFT & Comprehensibility	86	**0.40**	3.96	**0.000159**
L3_YFT & Correctness	86	**0.36**	3.56	**0.000617**
L3_AOL & Accent	86	**-0.49**	-5.09	**0.000002**
L3_AOL & Comprehensibility	86	**-0.36**	-3.52	**0.000690**
L3_AOL & Correctness	86	**-0.34**	-3.32	**0.001346**
N_TOTAL_Ln & Accent	86	-0.10	-0.89	0.376148
N_TOTAL_Ln & Comprehensibility	86	-0.06	-0.59	0.556408
N_TOTAL_Ln & Correctness	86	-0.02	-0.14	0.887054
Stay in L3 & Accent	28	0.10	0.53	0.603099
Stay in L3 & Comprehensibility	28	-0.03	-0.15	0.882157
Stay in L3 & Correctness	28	0.03	0.17	0.863867
Age & Accent	86	**0.22**	2.05	**0.043163**
Age & Comprehensibility	86	0.19	1.74	0.084747
Age & Correctness	86	0.10	0.95	0.346158
Eval L3 comp & Accent	86	**0.56**	6.16	**0.000000**
Eval L3 comp & Comprehensibility	86	**0.47**	4.91	**0.000004**
Eval L3 comp & Correctness	86	**0.43**	4.36	**0.000037**
Eval L3 pron & Accent	86	**0.54**	5.92	**0.000000**
Eval L3 pron & Comprehensibility	86	**0.48**	4.99	**0.000003**
Eval L3 pron & Correctness	86	**0.46**	4.79	**0.000007**
L2_Prof & Accent	86	0.20	1.88	0.063474
L2_Prof & Comprehensibility	86	**0.23**	2.20	**0.030773**
L2_Prof & Correctness	86	**0.26**	2.44	**0.016819**
L3_Prof & Accent	86	**0.66**	8.13	**0.000000**
L3_Prof & Comprehensibility	86	**0.60**	6.82	**0.000000**
L3_Prof & Correctness	86	**0.57**	6.28	**0.000000**
Year & Accent	86	0.06	0.51	0.611082
Year & Comprehensibility	86	0.05	0.43	0.667517
Year & Correctness	86	0.04	0.39	0.695507

4.6.3. Between variable correlations – joint analysis

Further Spearman rank correlations were performed to investigate interdependencies between pairs of speakers' variables. The present analysis was performed jointly on all the data from the four language groups under investigation. Significant correlations ($p<0.05$) were found to hold for several pairs of variables (see Table 38).

The length of formal training in L2 correlated strongly with the onset of L2 learning (negative correlation $R_S=-0.8$) and weakly with the onset of L3 learning ($R_S=0.27$). There was a weak negative correlation between the length of training in L2 and that in L3 ($R_S=-0.2$).

The speakers' age correlated moderately with their L2 proficiency level ($R_S=0.42$), L3 proficiency ($R_S=0.34$) as well as their year of studies ($R_S=0.59$), i.e. older participants were more proficient in their respective L2 and L3 and they were students of higher years.

Negative moderate to weak correlations were found between the age of onset of learning L3 and L3 proficiency ($R_S=-0.56$), overall self-evaluation of L3 general competence ($R_S=-0.32$) and self-evaluation of their own L3 pronunciation ($R_S=-0.29$), i.e. the later they started acquiring the L3, the lower their self-evaluation as well as their actual L3 proficiency.

The length of formal training in L3 correlated moderately with the participants' self-evaluation of L3 general competence ($R_S=0.43$) and self-evaluation of their own L3 pronunciation ($R_S=0.39$), and weakly with the year of studies ($R_S=0.23$). Strong correlations were reported between the length of L3 training and L3 proficiency level ($R_S=0.68$), and onset of L3 learning (negative correlation $R_S=-0.87$).

L3 proficiency correlated strongly with self-evaluation of L3 general competence ($R_S=0.7$) as well as with self-evaluation of their own L3 pronunciation ($R_S=58$), whereas the L2 proficiency level correlated moderately with the year of studies ($R_S=0.43$). Other correlations between the participants' variables were either weak or non-significant.

Table 38. Spearman rank correlations between pairs of speakers' variables – joint analysis

Pairs of variables	N	R_S	t(N-2)	p
L2_YFT & L2_AOL	127	**-0.80**	-14.98	**0.000***
L2_YFT & L3_YFT	127	**-0.22**	-2.57	**0.011***

L2_YFT & L3_AOL	127	**0.27**	3.11	**0.002***
L2_YFT & N_TOTAL_Ln	127	0.13	1.42	0.159
L2_YFT & Stay in L3	41	0.10	0.61	0.546
L2_YFT & Age	127	0.05	0.55	0.580
L2_YFT & eval L3 comp	127	**-0.35**	-4.11	**0.000***
L2_YFT & eval L3 pron	127	-0.15	-1.66	0.100
L2_YFT & L2_Prof	127	**0.39**	4.80	**0.000***
L2_YFT & L3_Prof	127	**-0.27**	-3.16	**0.002***
L2_YFT & Year	127	0.16	1.83	0.069
L2_AOL & L3_YFT	127	**0.43**	5.32	**0.000***
L2_AOL & L3_AOL	127	**-0.33**	-3.86	**0.000***
L2_AOL & N_TOTAL_Ln	127	-0.15	-1.64	0.104
L2_AOL & Stay in L3	41	0.07	0.44	0.661
L2_AOL & Age	127	**0.42**	5.12	**0.000***
L2_AOL & eval L3 comp	127	**0.45**	5.65	**0.000***
L2_AOL & eval L3 pron	127	**0.19**	2.15	**0.033***
L2_AOL & L2_Prof	127	-0.13	-1.51	0.134
L2_AOL & L3_Prof	127	**0.46**	5.76	**0.000***
L2_AOL & Year	127	0.14	1.52	0.130
L3_YFT & L3_AOL	127	**-0.87**	-20.12	**0.000***
L3_YFT & N_TOTAL_Ln	127	**-0.19**	-2.15	**0.033***
L3_YFT & Stay in L3	41	0.24	1.52	0.137
L3_YFT & Age	127	**0.40**	4.84	**0.000***
L3_YFT & eval L3 comp	127	**0.43**	5.37	**0.000***
L3_YFT & eval L3 pron	127	**0.39**	4.76	**0.000***
L3_YFT & L2_Prof	127	0.04	0.43	0.669
L3_YFT & L3_Prof	127	**0.68**	10.44	**0.000***
L3_YFT & Year	127	**0.23**	2.68	**0.008***
L3_AOL & N_TOTAL_Ln	127	0.15	1.64	0.104
L3_AOL & Stay in L3	41	-0.06	-0.35	0.725
L3_AOL & Age	127	-0.16	-1.86	0.065
L3_AOL & eval L3 comp	127	**-0.32**	-3.80	**0.000***
L3_AOL & eval L3 pron	127	**-0.29**	-3.37	**0.001***

L3_AOL & L2_Prof	127	0.10	1.14	0.257
L3_AOL & L3_Prof	127	**-0.56**	-7.61	**0.000***
L3_AOL & Year	127	-0.11	-1.22	0.225
N_TOTAL_Ln & Stay in L3	41	0.25	1.61	0.117
N_TOTAL_Ln & Age	127	-0.02	-0.26	0.799
N_TOTAL_Ln & eval L3 comp	127	-0.01	-0.11	0.910
N_TOTAL_Ln & eval L3 pron	127	-0.07	-0.82	0.414
N_TOTAL_Ln & L2_Prof	127	**0.21**	2.35	**0.020***
N_TOTAL_Ln & L3_Prof	127	-0.10	-1.15	0.253
N_TOTAL_Ln & Year	127	0.13	1.50	0.135
Stay in L3 & Age	41	0.07	0.41	0.684
Stay in L3 & eval L3 comp	41	0.18	1.11	0.272
Stay in L3 & eval L3 pron	41	0.06	0.40	0.688
Stay in L3 & L2_Prof	41	0.10	0.62	0.540
Stay in L3 & L3_Prof	41	0.17	1.11	0.274
Stay in L3 & Year	41	-0.09	-0.58	0.565
Age & eval L3 comp	127	**0.19**	2.21	**0.029***
Age & eval L3 pron	127	0.10	1.16	0.249
Age & L2_Prof	127	**0.42**	5.16	**0.000***
Age & L3_Prof	127	**0.34**	4.07	**0.000***
Age & Year	127	**0.59**	8.16	**0.000***
eval L3 comp & eval L3 pron	127	**0.67**	10.16	**0.000***
eval L3 comp & L2_Prof	127	-0.12	-1.34	0.182
eval L3 comp & L3_Prof	127	**0.70**	11.08	**0.000***
eval L3 comp & Year	127	0.04	0.49	0.622
eval L3 pron & L2_Prof	127	-0.04	-0.42	0.673
eval L3 pron & L3_Prof	127	**0.58**	8.04	**0.000***
eval L3 pron & Year	127	0.02	0.19	0.851
L2_Prof & L3_Prof	127	0.06	0.64	0.525
L2_Prof & Year	127	**0.43**	5.35	**0.000***
L3_Prof & Year	127	0.14	1.62	0.108

4.7. Discussion

The present section aims to discuss the results presented in the separate group sections by relating them to other findings from previous studies and interpreting them in the light of the theoretical underpinnings of the models proposed for third language acquisition. The purpose of this part of the study was to investigate the issue of foreign accentedness and possible sources of cross-linguistic influence in third language phonology. To this end, the research questions specified in the methodology section will be addressed.

RQ 1: How are the different rating parameters of the degree of foreign accent, comprehensibility and pronunciation correctness related in the perceptual judgements performed on L3?

The first research question investigated how different rating parameters of the degree of foreign accent, comprehensibility and pronunciation correctness are related in perceptual judgements performed on L3 samples. Interesting patterns of correspondence emerged between the L3 rating scores on 3 parameters that demonstrated a high degree of consistency across the four groups (A, B, C, D). The lowest mean scores were assigned for the degree of foreign accent in L3, with a total mean M=3.1 on a 7-point scale, ranging from 2.9 in Group C to 3.5 in Group D, with groups A and B scoring equally 3.1. The scores for comprehensibility were on average the highest, with a total mean M=4.9, ranging from 4.5 in C to 5.1 in D, and groups A and B having equal mean scores of 4.9. The mean pronunciation correctness scores turned out to be intermediate between the other two rating parameters, with a total mean score of M=3.8, ranging from 3.3 in Group C to 4.0 in Groups A and B. Furthermore, the conducted pairwise comparisons in each group point to statistically significant differences between the scores for each parameter, thus indicating that the rating parameters could not be conflated as the judges differentiated between three separate concepts for their perceptual assessment of L3 samples. Interestingly, the mean ratings for all the parameters were identical for Groups A with L3 French and B with L3 German although they were performed on different third languages. On the other hand, the ratings of L3 English for Groups C (with L2 German) and D (with L2 French) showed more variability, with Group C receiving somewhat lower scores.

A conducted joint comparative analysis across the four groups demonstrated no significant differences in the rating scores for particular parameters, with the exception of the correctness measure – in which Group C received significantly lower scores than Group A. All the rating parameters demonstrated high to moderate covariance with one another. The performed correlational analyses pointed to particularly high correlations between foreign accentedness and pronunciation correctness across the four groups, followed by consistently high correlations between comprehensibility and correctness ratings, as well as moderate correlations between the parameters of accentedness and comprehensibility.

The accent rating patterns observed in the present study show a considerable level of consistency with the author's previous related studies (cf. Wrembel 2010, 2012) with respect to the two rating parameters of foreign accentedness and comprehensibility, with the remaining parameters being modified in the present series of studies. Foreign accentedness attracted the lowest ratings, whereas comprehensibility generated much higher scores, thus substantiating the claim that a high degree of intelligibility is possible even though non-native speech is perceived as accented. The parameter of the rater's certainty level, used in the preliminary piloting studies, was excluded from the present series of investigations as it generated close to maximum scores and very little variability in the previous studies. Moreover, the sociophonetically-oriented parameter of acceptability from the previous explanatory investigation was replaced in the present study by the pronunciation correctness rating. It was intended to be a more objective and less attitudinally biased measure of pronunciation performance.

RQ 2: Do perceptual ratings in TLA exhibit similar patterns as those reported in SLA?

The second research question aimed to explore whether the perceptual ratings performed on a third language exhibit similar patterns to the ones reported in SLA. On the whole, the present series of accent ratings of L3 speech samples proved to be in line with the results reported in the SLA literature (e.g., Flege 1988; Gallardo del Puerto et al. 2007; Piske et al.

2001), according to which the degree of foreign accent is the most severely judged rating measure and where a lower degree of foreign accent is usually associated with higher intelligibility scores. The present findings provide further support for the claim put forward in SLA research that non-native speech may remain to a large extent intelligible in spite of some degree of foreign accentedness.

The observed variability in the accent ratings proved to be proficiency-related, as two measures of L3 proficiency (i.e. class assignment based on a competence test as well as self-assessment of the third language proficiency level) were found to correlate highly with all the rating parameters. As expected, the participants with higher proficiency in L3 scored better on the degree of foreign accent, comprehensibility and pronunciation correctness than their lower level counterparts. These results provide further validity to the study as the performed foreign accent ratings proved consistent with the subjects' actual proficiency levels. Moreover, this trend demonstrates an overall consistency with the findings of the L2 accent studies reported in the SLA literature (cf. Flege 1988; Gallardo del Puerto et al. 2007; Piske et al. 2001).

As far as inter-rater reliability is concerned, the statistical analysis performed by means of Cronbach alpha pointed to very high coefficients of internal consistency for the three rating parameters of accentedness, comprehensibility and correctness between raters across the groups (ranging from 0.98 – 0.94) as well as high coefficients for L1 identifications (in the range of 0.88 -0.76). Such high levels of interrater consistency are in line with the standard accent ratings in SLA studies and thus provide further evidence for the external validity of the study.

RQ 3: Do L1 identification patterns point to stronger native or non-native language influence on the perceived foreign accent in L3?

The third research question addressed the problem of sources of cross-linguistic influence in L3 phonological acquisition, i.e. whether it is the native or non-native language that causes stronger interference on the perceived foreign accent in L3. For the purpose of the present study, the cross-linguistic influence was operationalised as referring to the raters' perception of accentedness and their identification of the subjects' first language on the basis of speech samples of the third language.

The results of the L1 identification task show that the majority of the raters were not able to identify correctly the speakers' first language on the basis of their L3 performance. The correct identification as L1 Polish was observed on average in 36% of the ratings across the groups, ranging from 29% in Group C to 41.6% in Group B. The analysis looked also at the identification scores as the respective second language, and the average percentage of L1 identification as L2 was 12.8% for all the groups, with the lowest score of 9.4% in Group A, and the highest 17.5% in Group C. The across the groups analysis indicated that L1 identification as L1 Polish ranked the highest in the percentage of scores, followed by L1 identification as L2, which outscored in all cases L1 identification as other languages from the provided list of options (i.e. French, Russian, English, Spanish, German, and others).

All the remaining choices with the exclusion of L1 and L2, were treated jointly for the purpose of the present analysis and categorised as 'other languages' since the percentages of individual languages was low and did not display any specific tendencies. The resulting category received in the joint analysis the highest percentage of L1 identifications, i.e., on average 51%, ranging from 54% in Groups A and C to 48% in Groups B and D. However, since it is a composite value of small percentages assigned for individual languages, it should be interpreted rather as an indication of incorrect L1 identification as "other languages".

Group comparison between A and B (i.e. with different L3s but sharing English as the L2) showed comparable levels of L1 identification as the L2 English for both groups (ca. 10%). In turn, the between group comparison for C and D (i.e. with the same L3 English but different L2s), indicated higher levels of L1 identification as L2 (14% for L2 French in Group D and 17% for L2 German in Group C). Noteworthy, the L1 identification as French or German was found to be related to the group profile, i.e. the L1 identification scores as German were higher for Group C with L2 German (17.5%) vs. Group D (5.7%), and likewise for L1 identification as French in which Group D with L2 French (14%) outscored Group C (10%).

These findings demonstrate complex patterns of interaction between the language systems of multilingual speakers and provide further evidence against the claim that the first language (L1) is the only source of transfer in third language acquisition. It appears that the percentage of L1 identifications as L2 was also noticeable. Since L1 identification was operationalised in this part of the study as an indicator of a source language for cross-linguistic influence, the generated findings point to a stronger influence of the native language, i.e. a prevailing L1-accented speech in L3, but the impact of the second language was also attested, thus resulting in some instances of perceived L2-accented performance in the L3.

The results of this part of the series of studies substantiate to a large extent the first hypothesised scenario, namely, that the native L1 Polish influence would override the non-native influence, resulting in an L1-accented performance in L3. This is in line with the traditional view according to which the first language constitutes the main source of phonological transfer due to the articulatory motor routines established in first language acquisition. Therefore, the L1, as the basic constraint in phonology, supersedes any potential influence from other non-native languages in subsequent language acquisition. Previous indications of a prevailing L1 influence on L3 phonology were reported in the multilingual literature e.g. by Ringbom (1987) and Pyun (2005).

The observed patterns of identification partially disconfirm the second hypothesised outcome that the non-native influence of respective L2s, i.e. the so called 'L2 status', is the prevalent source language for cross-linguistic influence leading to a perceived L2-accented speech in L3. Consequently, the study fails to provide strong support for the L2 status factor model (Bardel & Falk 2007; Falk & Bardel 2011) and the primacy of the non-native influence as reported in some L3 studies (e.g. Hammamberg & Hammamberg's 1993, 2005; Llama et al. 2010), although the participants' proficiency in this study was higher.

Nevertheless, the present findings provide partial support for the third hypothesised general outcome, according to which both the native and non-native languages exert an impact on the perceived foreign accent in the L3, thus verifying the assumption of a combined cross-linguistic influence (cf. de Angelis 2007). Such an interpretation of the data lends some validity to the Cumulative-Enhancement Model for Language Acquisition (Flynn et al. 2004), according to which all the previously learnt languages may influence the acquisition of the third language under the condition that the transfer is facilitative.

RQ 4: Which factors, including raters' variables, have any bearing on the results?

The fourth research question focused on the factors which could have a potential bearing on the results; including the speakers' variables and the raters' variables.

As far as the speakers' characteristics are concerned, the analysis accounted for such factors as their L3 proficiency level, years of formal training in L3, participants' age, age of onset of L3 learning, self-evaluation of L3 proficiency, year of studies, stay in an L3 speaking country or the number of foreign languages known by the participants. The groups showed two different patterns of correlation, with Groups A and B featuring several strong to moderate correlations, whereas Groups C and D showed only a moderate correlation with a single variable. The speakers' variables that were correlated with all the rating parameters included L3 proficiency, years of formal learning in L3, age of onset of L3 learning, self-evaluation of L3 proficiency in Groups A and B as well as the participants' age and year of studies additionally in Group A. On the other hand, the ratings correlated moderately only with age of onset of L3 learning in Group C and the year of studies in Group D.

A detailed analysis of the observed correlations indicated that the more proficient the participants were in their respective L3s, the more years of formal training they had had, the higher their self-evaluation of L3 proficiency and the earlier they started learning their L3, the better they were rated on accent, comprehensibility and pronunciation correctness. The directionality and strength of the reported covariance confirms the expected trends, verifying at the same time the internal validity of the performed ratings. The lower number of significant correlations generated for Groups C and D, which were subject to joint ratings of L3 English, could be the result of a smaller statistical power of the tests performed separately for these two subcomponent groups, based on a more limited number of speech samples than in the case of the rating analyses for Groups A and B. A joint correlational analysis for all the groups, which featured a stronger statistical power than the individual group analyses, indicated that the L3 proficiency level as well as the years of formal training in L3 are the strongest predictors of perceptual L3 pronunciation performance as evaluated by the means of accent ratings.

As far as the raters' variables are concerned, interesting interdependency trends were reported across the groups. Groups A and B exhibited similar patterns of correlations between the raters' variables and accentedness ratings in L3. In the case of the native vs. non-native status effect, significant differences were found for one out of three rating parameters; in Group A native speaker raters assigned lower scores for foreign accent, whereas in Group B native raters gave lower scores for comprehensibility. The difference was not found to be significant in Groups C and D. The trends were identical for the raters' proficiency level variable.

These findings do not substantiate previous research results (e.g. Scheuer 2000; Wrembel 2010, 2012), which indicated that native raters tend to be more lenient in their foreign accentedness ratings, whereas some of the present data point to an opposite trend for individual rating parameters. On the whole, for the most part of the analysis, native and non-native raters do not exhibit any significant differences in the performed L3 pronunciation assessment. These results lend further support to the experimenter's decision to include non-native raters as legitimate judges able to perform the rating tasks in a comparable manner to the native speakers.

As far as L1 identification scores are concerned, the raters' nativeness variable proved more relevant as significant differences were reported in all the groups. L1 identifications as L1 Polish were performed much better by non-native raters who were actually Polish native speakers proficient in the respective foreign languages. On the other hand, native raters (i.e. English, French or German native speakers in each group respectively) tended to identify the speakers' L1 more frequently incorrectly as other foreign languages. The observed trends are consistent with the expectations that the Polish raters would find it easier to identify correctly the mother tongue of Polish foreign language learners on the basis of their performance in the L3 on account of the shared L1. Apparently, this group of raters was able to identify better the traces of L1 accent in L3 speech produced by their compatriots.

The analysis of the impact of raters' phonetic training on the performed ratings pointed to statistically significant differences for one or two rating parameters for three out of the four groups. The phonetically trained raters tended to assign lower scores for pronunciation correctness in Group A; lower scores for pronunciation correctness and comprehensibility in Group B; and higher scores for comprehensibility in Group D. The variable of phonetic training did not generate any statistically significant differences in Group C. Summing up, mixed results were reported as far as the effect of phonetic training is concerned; however, this variable appeared to exert some impact on the evaluation of phonetic performance.

Raters' phonetic training did not affect their L1 identification patterns in a uniform manner either. The findings were mixed in this respect, with a tendency for phonetically trained raters to identify better the speakers' L1. On the other hand, phonetically untrained raters tended to identify the speakers' L1 as their respective L2 more frequently than trained raters.

All in all, the raters' variables did not generate consistently different patterns across the groups apart from the observed tendencies discussed above. It appears that the raters' differences concerning their native status, level of proficiency and phonetic training cannot be regarded as very strong predictors of the L3 pronunciation perceptual evaluation, but rather as indications of some trends that need further verification.

Because of the complexity of the present research design, only read samples were rated, therefore no direct comparisons to the author's previous results (e.g. Wrembel 2010, 2012b) evidencing the effect of the L3 performance mode (i.e. read vs. spoken speech) for the identification of speakers' L1 are possible. However, the prevalence of L1 identifications as L1 Polish in this series of studies can be, to some extent, also attributed to the more controlled reading mode of L3 performance, which according to previous results, exhibits more phonetic features transferred from the first language. Conversely, the subjects' spoken L3 samples were found to be more L2-accented (cf. Wrembel 2010, 2012b). According to the author, this less monitored performance mode demonstrates more traces of the second language resulting in higher L1 identification as L2, thus substantiating the hypothesised unconscious switch to a 'foreign language mode' (cf. Cohen 1995) that is less evident in the more controlled reading mode.

As far as the source of cross-linguistic influence is concerned, the present results of the L1 identification task in accent ratings indicate that it is more the native tongue rather than non-native language that acts as the source language in L3 phonological acquisition. The present data provide more support for the claim that it is the motor routines of the mother tongue that dominate the acquisition of the third language phonology rather than a previously learnt other foreign language (cf. Gut 2010; Ringbom 1987). There are some indications, however, based on the patterns of L1 identification as the second language, that L2 exerted some impact on the L3 interphonology, yet the strength of this influence did not prove to be as significant as in the case of L1.

The present findings from a series of accentedness ratings performed on L3 French, L3 German and L3 English on the participants of Groups A, B, C and D complement previous exploratory investigations into L1 Polish/ L2 English/ L3 French vs. L1 Polish/ L2 French/ L3 English mirror-design pairings (Wrembel 2012a, b). On the whole, the reported tendencies proved to be comparable, i.e. the L3 speech was mostly found to be L1-accented with some evidence of a complementary L2 effect. Different evidence was generated, however, in the earliest investigation of this kind on L1 Polish/ L2 German/ L3 English (Wrembel 2010), where the L3 samples were found to be prevailingly L2-accented with an underlying L1 effect. Interestingly, in the present series of studies, it was Group C with an identical language repertoire, i.e. L1 Polish/ L2 German/ L3 English, that exhibited the highest percentage of L2-accentedness based on L1 identification rates compared to other groups (A, B and D).

As the perceived non-native accentedness in the third language performance was much more salient in the case of typologically related pairs of languages, i.e. German and English, rather than the less related English and French pairings, the author concluded that it is the additional factor of language distance rather than foreign language effect (or ‘L2 status’) alone that may condition the source and extent of a cross-linguistic influence in L3 phonological acquisition. The author’s findings may be interpreted as being consistent with Rothman’s (2011, 2015) Typological Primacy Model (TPM), according to which the selection of the native vs. non-native source of transfer in the acquisition of a third language is determined by the typological proximity between the target language and L1 or L2, as perceived by the learner’s internal mechanisms, rather than by the order of acquisition alone.

It stems from the assumptions of the TPM that the unconscious internal mechanisms operating in L3 acquisition are aimed at transferring knowledge from previously acquired languages. This process relies on an internal diagnostic set determining the optimal choice for similarity and it has economic motivations. The languages under investigation (i.e. Polish, French, English and German) exhibit different patterns of typological relatedness and some predictions can be made about the L3 transfer mechanisms on the basis of the phonetic and phonological proximity. The features that are particularly significant for oral comprehension and, at the same time, for perceived pronunciation assessment, involve temporal patterns (cf. Cutler et al. 1986). The set of languages involved in the accent ratings exhibit some major differences in their prosodic structure; with English and German being stress-timed languages, French, on the other hand, being syllable-timed, and Polish classified either as mixed or even stress-timed, in spite of the lack of evident vowel reduction (cf. Grabe & Low 2002). Therefore, based on temporal patterns, English and German, as well as, to some extent, Polish, can be expected to exhibit closer typological proximity. On the other hand, the predictions related to the phonemic structure may indicate more similarity between French, English, and German as these languages have much larger vocalic repertoires than Polish and exhibit vowel lengthening in certain consonantal contexts. Consequently, the predictions are mixed depending on the selected criteria for typological proximity, which may account for the rather complex pattern of results evidenced in the performed series of L3 accent ratings.

Further, it may be hypothesised that yet another factor may account for the complexity of the results, namely, the actual L2 proficiency level. If we assume that a prerequisite for L2-accented speech in L3 is a sufficiently high level of L2 proficiency rather than the initial stage of L3 acquisition alone, as proposed e.g. by Hammarberg & Hammarberg (1993, 2005) or Gut (2010), such an interpretation could partially explain the finding/s patterns in the present series of accent studies. In Wrembel's (2010) study the subjects' L2 German proficiency level was advanced and generated a high percentage of perceived L2-accented speech in L3 English ratings, whereas in the present series of studies, for instance, in Group A, the participants' L2 French was at an upper-intermediate level, which presumably was not sufficient to result in a strongly perceived non-native accent in their L3 English.

The present findings partially disconfirm Hammarberg and Hammarberg's (1993, 2005) claim that it is the other non-native language

(L2) rather than the mother tongue that constitutes the stronger source language in L3 phonological acquisition thus resulting in an L2-accented speech in L3 performance. The present results, based on foreign accent ratings performed on L3 French, L3 German and L3 English, lead to the conclusion that it is the native language that is prevalent as the source of cross-linguistic influence in the phonological acquisition of L3; however, the influence of the non-native tongue seems to be also noticeable.

On the whole, the results appear to be consistent with the assumption of a combined cross-linguistic influence that involves the simultaneous influence of more than one previously acquired languages on the target language (cf. De Angelis, 2007). This claim, posited from the perspective of Third Language Acquisition, constitutes an extension of a traditional SLA view of cross-linguistic influence (CLI), which has been perceived to be of a one-to-one type between the source and the target language. The present study substantiates the existence of a combined CLI, although not to the extent evidenced in some other L3 studies.

Chapter 5

Study II: VOT patterns in L3 acquisition

5.1. Research design

This part of the study investigates a selected aspect of foreign accentedness by focusing on the acquisition of laryngeal contrasts (i.e. voice onset time) in a third language. It is meant as an extension and large scale verification of previous preliminary studies on VOT patterns in third language acquisition conducted by the present author on L3 French and L3 German (cf Wrembel 2011, 2014, 2015). The present study widens the perspective by comparing the VOT acquisition patterns across various language combinations in four groups (A, B, C, D).

5.2. Study aims and research questions

The main goal of this series of studies was to explore the complexity of voice onset time (VOT) patterns in trilingual acquisition. More specifically, it aimed to investigate in detail the sources, directionality and relative strength of cross-linguistic influence (CLI) in third language acquisition focusing on this selected acoustic parameter. Further, four studies (A, B, C, D) were designed in parallel to provide the ground for a comparison of the acquisition patterns of laryngeal contrasts in various combinations of the selected language repertoires.

In order to address the objectives specified above, the following research questions were posed:

RQ 1: Do multilingual subjects distinguish between their language systems (i.e., L1, L2 and L3) with respect to the VOT values?
RQ 2: Do the L3 VOT patterns approximate the participants' values in their L1, L2 or the L3 native norms?
RQ 3: Is there a proficiency effect on the L3 VOT measurements based on the amount of L3 exposure?
RQ 4: Which other factors have an impact on the VOT values in the three languages?

RQ 5: Does the typological proximity between particular language repertoires influence the reported VOT patterns?
RQ 6: Do VOT patterns in trilingual acquisition resemble those reported in SLA literature or are these trends specific for each context of acquisition?

On the basis of the overview of current literature on third language acquisition, four potential general outcomes were hypothesised:

(1) native L1 Polish would prevail as the source of cross-linguistic influence for the acquisition of VOT patterns in the L3;
(2) the so called ‘foreign language effect’, i.e. the influence of the participants’ respective L2 would override the native language in shaping L3 VOT values;
(3) both the native and non-native languages would have an impact on the VOT values in the L3, thus substantiating the assumption of a combined cross-linguistic influence;
(4) the source of the CLI reflected in VOT measures would be determined by the typological proximity between particular language systems rather than the L1 or L2 language status.

5.3. Procedure

The data collection procedure involved all three language systems of the multilingual participants, i.e. their respective L1, L2 and L3. All the participants of Groups A, B, C, and D took part in Study I (see Chapter 4, section 4.1.3 for a detailed description of the participants' profiles). Out of the total number of 128 participants, 3 had to be excluded from the subsequent analysis due to incomplete data collection in one of their language systems. Therefore, the VOT measures were recorded in the first, second and third languagesof the following number of participants; Group A N=38, Group B N=26, Group C N=33, Group D N=28.

The stimuli consisted of three word lists with 18 target words and 12 distractors in each respective language. The target words included voiceless plosives /p, t, k/ in stressed onset positions in the following context of high vs. mid and low vowels, in mono- and disyllabic words, thus generating a total of 18 items per language list. The words were randomized and embedded in carrier phrases in particular languages (i.e., I am saying ..., Mówię teraz …, Je dis …, Ich sage …).

The recordings were made in a clearly specified language mode in the natural order of the acquisition of the languages involved, with the L1 as first, the L2 as second and the L3 third. There were breaks between different language sessions and the interaction with the researcher was carried out in the language of the subsequent recording to promote the activation of the respective languages. To ensure a controlled speed of delivery, as reading pace is an important factor influencing VOT values, the stimuli in the carrier phrases appeared on the computer screen at regular intervals of time having been incorporated into a PowerPoint presentation.

The stimuli were recorded in an attenuated booth in a recording studio with the application of Audition CS5.5 as 16-bit mono files at a 32 000Hz sampling frequency. All the tokens were subjected to an acoustic analysis performed using PRAAT 5.2.15 (Boersma and Weenick, 2010). Tokens were excluded from the analysis if the target words were mispronounced. Voice onset time was measured in milliseconds (ms) as the interval between the release burst and the beginning of the regular vocal fold vibrations. After the recording session, the participants were requested to fill in a language background questionnaire to tap their language history and use.

The languages selected for this series of studies, including Polish, French, German and English, make a phonological distinction between two categories of stops; however, their phonetic realisation differs. While English and German belong to the category of the so called aspirating languages (cf. Lisker and Abramson, 1964), which differentiate between voiceless aspirated and voiceless unaspirated plosives, Polish and French are voicing languages, which make a distinction between voiced and voiceless unaspirated plosives. Therefore, in Polish and French /p/, /t/, /k/ are implemented as short-lag stops, with mean VOT values around 20-50 ms for Polish (Keating et al., 1981), and 20-30 ms for French (Caramazza et al., 1973). On the other hand, in English and German /p/, /t/, /k/ are implemented as long-lag stops with average VOT values for English at around 60-80 ms (Lisker and Abramson, 1964), and for German between 30-50 ms (Angelowa and Pompino-Marschall, 1985).

5.4. Results analysis

The conducted analysis of the results was based on the acoustic measurements of the voice onset time of the target words read in the carrier phrases in all three languages of the multilingual participants, i.e., their L1 Polish, L2 German and L3 English. The statistical analyses were performed using the Statistica software. They involved (1) mean VOT values for the L1, L2 and L3, (2) cross-language comparison of VOT means, (3) proficiency group effect analysis, (4) comparison to VOT literature reference values, and to control groups' values, (5) the analysis of the vocalic context effects and (6) the analysis of variance and correlation analysis accounting for the relationships between independent variables. The results of the subsequent analyses will be presented in the following sections.

5.4.1. Results for Group A

In the following sections, the results of the analysis for Group A with L1 Polish, L2 English and L3 French will be presented in detail.

5.4.1.1. Mean VOT values for L1, L2 and L3

First, the Shapiro–Wilk test of normality was administered to verify the null hypothesis for this test that the VOT measurements data are normally distributed ($p>\alpha=0,05$). Since the chosen alpha level was 0.05 and the p-values for the VOT variable in several categories were less than 0.05, then the null hypothesis that the data are normally distributed had to be rejected. Consequently, non-parametric tests were applied for the further statistical analyses of the data in Group A.

Tables 39-41 present the mean results of the VOT measurements of the voiceless plosives /p/, /t/, /k/ of the target words read in the carrier phrases in the participants' L1 Polish, L2 English, and L3 French.

Table 39. Mean VOT values for target words in L1 Polish

Words in L1	N	Mean	SD	Median
pan	38	20.4	10.5	18.0
para	38	18.9	9.8	16.0
pech	38	16.6	6.4	15.5
pole	38	25.3	12.1	21.5
pik	38	24.6	10.1	23.5
piwo	38	30.9	13.2	28.5
tak	38	21.7	7.4	21.5
tango	38	24.7	9.0	22.0
tor	38	26.3	9.5	25.0
testy	38	23.6	7.4	23.5
tik	38	38.4	12.0	36.0
tiry	38	44.1	18.1	42.5
kat	38	46.7	12.2	45.0
kanon	38	44.6	12.7	41.0
kot	38	46.9	10.9	45.5
kelner	38	41.8	10.9	40.0
kit	38	59.1	12.4	57.5
kino	38	63.3	16.1	67.0

Table 40. Mean VOT values for target words in L2 English

Words in L2	N	Mean	SD	Median
pan	38	80.6	32.4	82.5
party	38	67.5	25.9	66.0
pot	38	64.4	26.5	58.5
pencil	38	48.6	25.7	43.0
peace	38	68.1	25.4	66.0
Peter	38	60.9	23.7	60.5
task	38	73.4	26.9	72.5
taxi	38	66.0	30.0	65.0
test	38	70.4	28.3	65.5
toffee	37	71.2	29.7	68.0
team	38	88.3	24.3	88.5
teacher	38	75.9	23.7	71.5
cat	38	98.2	23.6	98.0
carpark	38	87.1	19.6	83.0
kept	38	82.7	18.5	82.5
coffee	38	74.0	18.2	71.5
keen	38	104.3	19.7	102.5
keeper	38	94.6	18.3	94.5

Table 41. Mean VOT values for target words in L3 French

Words in L3	N	Mean	SD	Median
pas	38	43.8	24.5	39.5
part	38	28.1	16.6	24.0
poche	38	40.6	18.8	36.0
perdre	37	29.0	18.4	23.0
pic	38	39.7	22.9	36.0
pile	38	55.2	20.4	51.5
taxe	38	35.0	20.1	30.5
table	37	41.4	20.3	34.0
terre	38	37.6	17.6	32.5
torse	38	34.4	15.3	32.5
tic	38	54.8	23.1	52.0
titre	38	50.7	15.6	49.0
calme	38	65.9	26.5	63.0
carte	38	56.2	19.5	53.0
coq	37	65.2	19.9	62.0
quelle	36	62.9	21.1	63.0
qui	31	81.6	15.3	81.0
quiche	28	75.8	16.9	74.5

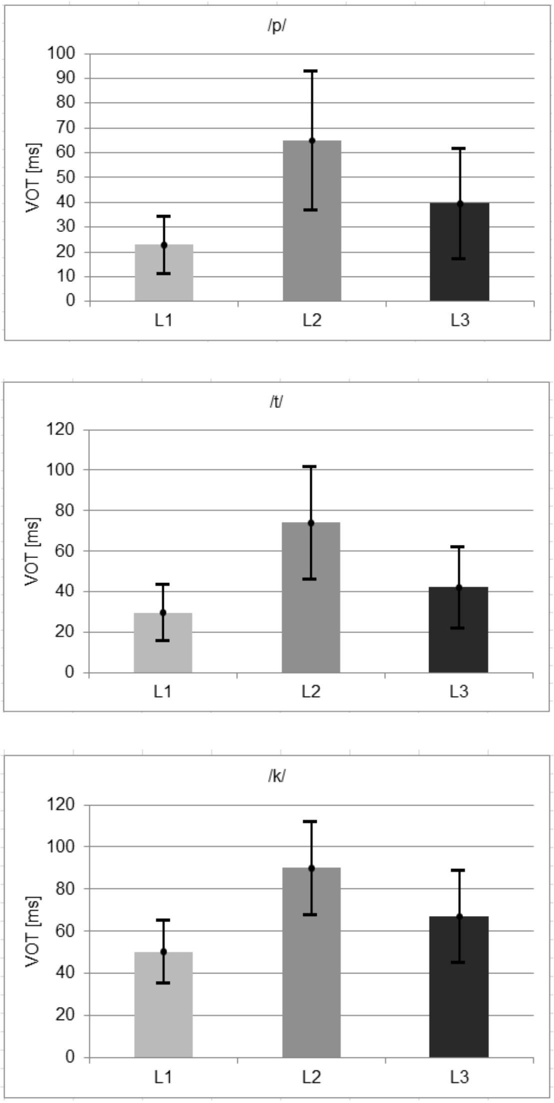
The acoustic measurements of the mean voice onset time for the voiceless plosives /p/, /t/, /k/ in stress onset positions were performed in the participants' L1 Polish, L2 English and L3 French. The results are presented in Table 42 and Figures 18-20.

The conducted durational measurements demonstrate that the multilingual participants produced voiceless plosives in the stress onset positions, with mean voice onset time values that were the shortest in L1 Polish (/p/=23 ms, /t/=30 ms, /k/=50 ms), had a typical long-lag in L2 English (/p/=65 ms, /t/=74 ms, /k/=90 ms), and demonstrated intermediate values in L3 French (/p/=39.5 ms, /t/=42 ms, /k/=67 ms).

Table 42. Mean VOT values for L1, L2 and L3

VOT	Language	N	Mean	SD	Median
/p/	L1	228	22.8	11.5	20.0
	L2	228	65.0	28.1	64.0
	L3	227	39.5	22.2	35.0
/t/	L1	228	29.8	13.9	27.5
	L2	227	74.2	27.8	72.0
	L3	227	42.3	20.2	36.0

/k/	L1	228	50.4	14.8	48.0
	L2	228	90.1	22.0	88.0
	L3	208	67.2	21.8	67.0



Figures 18-20. VOT measurements for /p, t, k/ in L1 Polish, L2 English, L3 French

5.4.1.2. Cross-language comparison of VOT means

A series of statistical tests was run with the view to investigating the language effect on the VOT durations in Group A. A non-parametric Kruskal-Wallis ANOVA was conducted in order to compare the mean VOT values for /p/, /t,/ /k/ across the participants' languages (L1, L2, L3) with the assumed alpha level being 0.05.

The results of ANOVA pointed to significantly different values for all the initial voiceless plosives between L1 Polish and L2 English, between L1 Polish and L3 French as well as between L2 English and L3 French ($p<.05$), see Table 43.

Table 43. Mean VOT comparison for /p t k/ in L1 Polish, L2 English, L3 French

Kruskal-Wallis Test	p for multiple comparisons		
	L1 vs L2	L1 vs L3	L2 vs L3
/p/	0.000000*	0.000000*	0.000000*
/t/	0.000000*	0.000000*	0.000000*
/k/	0.000000*	0.000000*	0.000000*

*p<0.05

As far as the cross-linguistic measurements are concerned, the observed patterns were identical for all the voiceless plosives /p, t, k/ and the p values for the multiple comparisons pointed to statistically significant differences in the mean VOT between all the language pairings:

– between L1 and L2 (mean VOT values in L1 Polish are lower than in L2 English)
– between L1 and L3 (mean VOT values in L1 Polish are lower than in L3 Frenc h)
– between L2 and L3 (mean VOT values in L2 English are higher than in L3 French).

All in all, the performed pairwise comparisons of the means showed cross-linguistic differences that proved to be significant between the native Polish and both of the non-native languages as well as between the L2 English and L3 French.

The following box plots (Figures 21-23) illustrate the tendencies observed in the VOT patterns in the respective languages separately for the stressed onset plosives /p/, /t/ and /k/. Particularly noteworthy is that the distribution of the VOT means is different for all the languages involved, whereas the standard deviation and the minimum-maximum range is wider for the non-native languages when compared to L1 Polish.

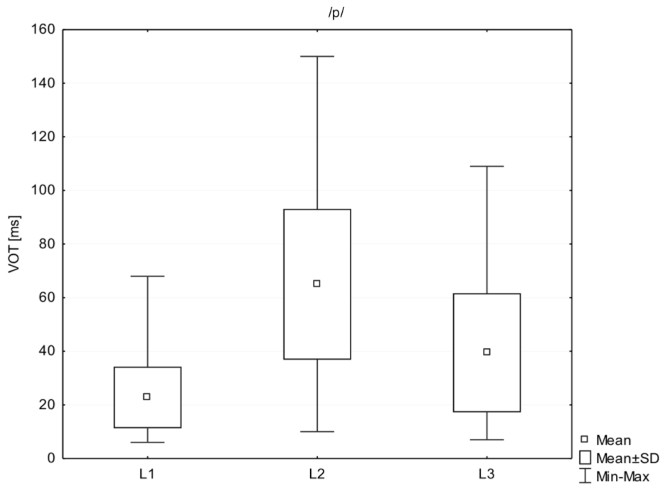


Figure 21. Box plot of mean VOT values for /p/ in L1 Polish, L2 English and L3 French

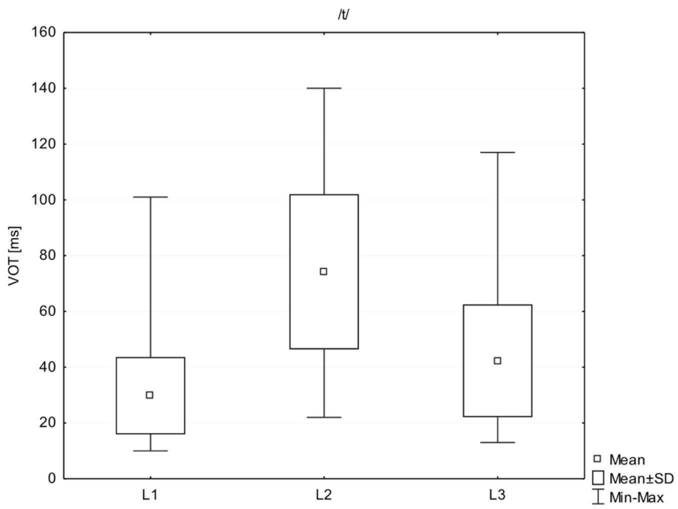


Figure 22. Box plot of mean VOT values for /t/ in L1 Polish, L2 English and L3 French

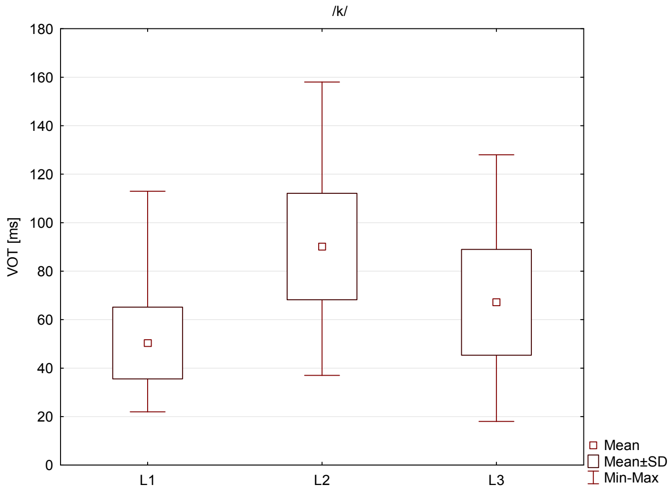


Figure 23. Box plot of mean VOT values for /k/ in L1 Polish, L2 English and L3 French

Recapitulating, the language effect was observed to occur between all of the phonological systems, including the native and non-native ones. The mean voice onset time values in L3 French were found to be intermediate between the L1 and L2 values, i.e. they were significantly higher than the respective values in L1 Polish and significantly lower than the L2 English VOT.

5.4.1.3. Proficiency group effect

The language proficiency group effect on the VOT acquisition in Group A was examined by means of non-parametric Mann-Whitney U tests, which were performed for the mean VOT values for /p/, /t/, /k/ in L1, L2 and L3 for the lower proficiency group 1F (N=18) vs. the higher proficiency group 2/3F (N=20) (see Table 44).

Table 44. Mann-Whitney test between groups 1F vs. 2/3F for /p, t, k/ in L1, L2 and L3

Sounds	Language	Z corr.	p
/p/	L1	0.05	0.961477
	L2	**-4.2**	**0.000027***
	L3	**2.04**	**0.041829***
/t/	L1	0.83	0.404334
	L2	**-4.5**	**0.000006***
	L3	**3.09**	**0.002033***

/k/	L1	1.34	0.179384
	L2	**-2.4**	**0.016162***
	L3	1.59	0.110721

*p<0.05

The results of the analysis demonstrated significant proficiency group effects with respect to the differences between the mean VOT for /p, t, k/ in L2 English and for /p, t/ in L3 French. No significant proficiency group effects were found in the case of L1 Polish.

The observed statistically significant differences between the 1F and 2/3F proficiency groups for the mean VOT values were as follows. For L2 English, the mean VOT values for /p/, /t/, /k/ were lower in the 1F group than in the more advanced 2/3F group. For L3 French, the opposite trend was true, i.e. the mean VOT values for /p/ and /t/ were statistically higher in the 1F group than in the 2/3F group. The recorded trends corresponded to the expected effects, as the higher proficiency group (2/3 F) exhibited values that were closer to the target range of the VOT values in English as well as in French (i.e. higher for L2 English and lower for L3 French).

The following box plots (Figures 24-26) illustrate the observed tendencies in the group comparison based on the L3 proficiency criteria (1F – less advanced, 2/3F – more advanced group).

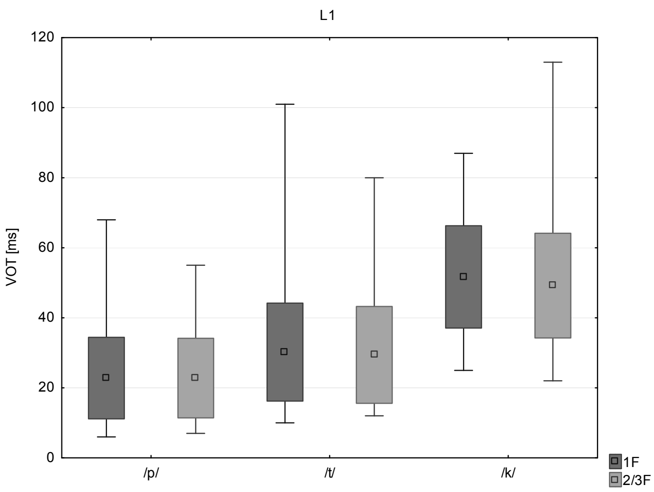


Figure 24. Box plot of mean VOT values for /p/, /t/, /k/ in L1 Polish – group comparison

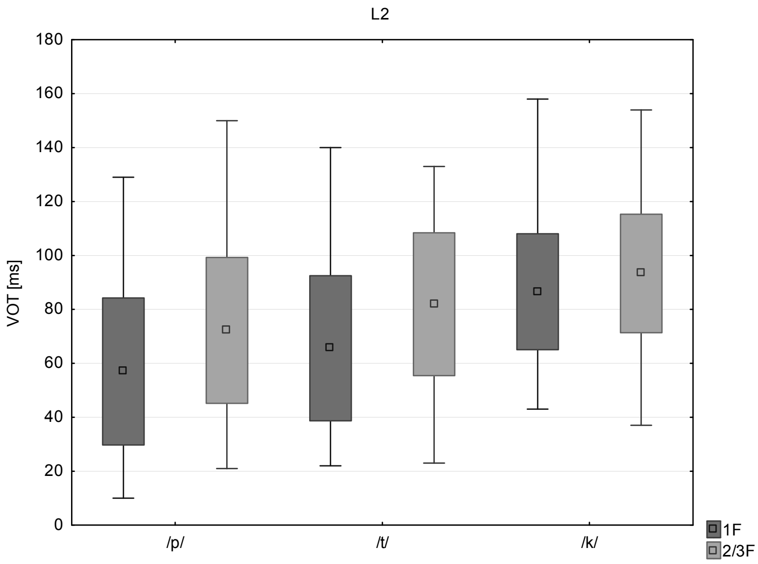


Figure 25. Box plot of mean VOT values for /p/, /t/, /k/ in L2 English – group comparison

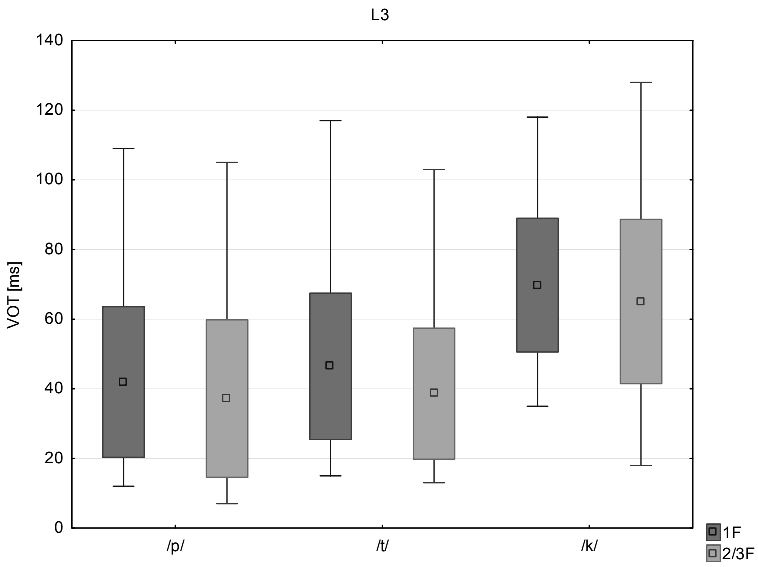


Figure 26. Box plot of mean VOT values for /p/, /t/, /k/ in L3 French – group comparison

5.4.1.4. Cross-linguistic correlations between VOT values

A Pearson product-moment correlation analysis was performed in order to investigate the correlation between two variables, i.e. the observed VOT values in L3 French and those for the native L1 Polish as well as between L3 French and L2 English (the assumed alpha level was α=0.05). The calculated coefficients pointed to prevailingly positive weak to moderate correlations between the mean VOT values in Group A.

For the voiceless bilabial plosive /p/, there were statistically significant correlations for both pairs of variables. The results pointed to a weak positive correlation (R=0.25) between L1 Polish and L3 French as well as a weak correlation (R=0.2) between both the non-native languages L2 English and L3 French in the mean VOT values for /p/ (see Table 45).

Table 45. Pearson's correlation for /p/

Pair of variables for /p/	N	R	t	p
VOT L1 & VOT L3	227	0.25	3.93	0.000111*
VOT L2 & VOT L3	227	0.20	3.13	0.001984*

*p<0.05

In the case of the alveolar plosive /t/, the obtained Pearson's correlations were significant for both pairs of variables; with a moderate correlation for L1 Polish and L3 French (R=0.35), and a weak one for L2 English and L3 French (R=0.24) (Table 46).

Table 46. Pearson's correlation for /t/

Pair of variables for /t/	N	R	t	p
VOT L1 & VOT L3	226	0.35	5.52	0.000000*
VOT L2 & VOT L3	226	0.24	3.68	0.000287*

*p<0.05

A similar pattern of correlations was observed for VOT in the velar plosive /k/, with a moderate positive correlation found to hold between L1 Polish and L3 French (R=0.4), and a weak one between L2 English and L3 French (R=0.29) (Table 47).

Table 47. Pearson's correlation for /k/

Pair of variables for /k/	N	R	t	p
VOT L1 & VOT L3	208	0.40	6.28	0.000000*
VOT L2 & VOT L3	208	0.29	4.35	0.000021*

*p<0.05

As the correlations between the VOT durations for both language pairings are significant in the case of all the plosives and they are rather similar in magnitude, only one selected scatterplot will be presented to illustrate the pair of variables with a higher correlation coefficient, i.e., L1 Polish and L3 French (see Figure 27). Recapitulating, the covariance between the respective variables remained in the weak to moderate range, with a tendency for slightly higher correlations between the native tongue L1 Polish and L3 French VOT values for all the plosives.

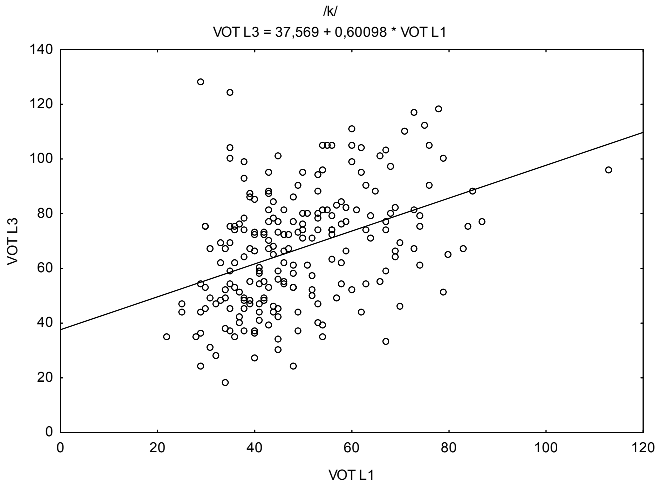


Figure 27. Scatterplot of VOT for /k/ between L1 Polish and L3 French

5.4.1.5. Comparison to VOT reference values

Further statistical analyses were performed, including one-sample t-tests and a non-parametric Wilcoxon signed-rank matched pairs test in order to compare the acoustic measurements obtained in Group A for /p, t, k/ in L1 Polish, L2 English and L3 French to the monolingual reference values, as quoted in the literature, i.e. Keating et al. 1981 for Polish; Lisker and Abramson 1964 for English and Caramazza et al. 1973 for French (see Table 48).

Table 48. Comparison to VOT reference values in Polish, English and French ([1] Keating et al., 1981; [2] Lisker and Abramson, 1964; [3] Caramazza et al. 1973)

Language	Parameter	VOT		
		/p/	/t/	/k/
Polish	Ref. VOT[1]	22	28	52
	L1 Mean	22.8	29.8	50.4
	SD	11.5	13.9	14.8
	p	0.274687	0.682606	0.109057
English	Ref. VOT[2]	59	67	84
	L2 Mean	65	74.2	90.1
	SD	28.1	27.8	22
	p	0.008770*	0.000795*	0.000193*
French	Ref. VOT[3]	18	23	32
	L3 Mean	39.5	42.3	67.2
	SD	22.2	20.2	21.8
	p	0.000000*	0.000000*	0.000000*

*p<0.05

The obtained measurements demonstrate that the VOT durations observed in the study for L2 English and L3 French differ significantly from the monolingual reference values, as reported in the literature, whereas the VOT values for L1 Polish correspond to the reported norms for Polish (see Table 48). In fact, the VOT values for the voiceless stops in L1 Polish were very close to the monolingual norms from the literature (cf. Keating et al., 1981), i.e. nearly identical in the case of /p/ and within 2 ms range for /t/ and /k/ (29.8 vs. 28 ms; 50.4 vs. 52 ms).

Both in L2 English and L3 French, the observed values differed significantly from the reference values; however, the directionality of the discrepancy varied. In the case of L2 English, the VOT values generated in the study exhibited a somewhat longer mean duration of voice onset time than the literature reference values (Lisker and Abramson, 1964); however, the difference was within 6-7 ms range, with 65 vs. 59 ms for /p/, 74 vs. 67 ms for /t/, and 90 vs. 84 ms for /k/. We can observe an overshoot of the target values when compared to the norms. On the other hand, in L3 French the onset voiceless plosives were realized with a considerably longer lag than the monolingual reference norms (Caramazza et al. 1973), with 39.5 vs. 18 ms for /p/, 42 vs. 23 ms for /t/, 67 vs. 32 ms for /k/. In the L2 English the observed VOT values approximated the target norms as reported in the literature, exhibiting an overshoot which was still within the accepted 5-10 ms range; the mean VOT measurements in L3 French demonstrated intermediate values between the L1 VOT and L2 English values. Particularly noteworthy is that the L3 French VOT values were significantly longer than the L1 Polish as well as the target native French durations, so there was no facilitative effect of the L1 transfer in this respect as could have been expected. An explanation for the longer lag found in the L3 French values may be a potential influence of the non-native well established VOT values in L2 English.

The native system of the participants does not appear to be affected by other non-native languages as the VOT durations for L1 Polish quite closely resembled those of the monolingual norms.

On the whole, the multilingual participants seem to be aware of the differences in the phonetic realisation of the laryngeal contrasts in the different languages of their repertoire and, thus, they tried to keep their language systems apart although the L3 French does not necessarily approximate the reference target but instead exhibits some intermediate, hybrid values.

5.4.1.6. Comparison to Control groups

With a view to verifying further the reference values found in the literature, the VOT measurements were generated from the native control participants involved in the study in Group A. To this end, a similar recording procedure was followed as in the case of the experimental group with the exception that the controls performed only in their native tongue. The French control group consisted of 8 native French lecturers at Adam Mickiewicz University in Poznań, who were recorded reading the word list in the carrier phrases for French. The English control group involved 17 native English speakers (i.e. lecturers at the University of Essex and Adam Mickiewicz University in Poznań), who were recorded reading the list for English. The generated control mean VOT values were /p/=29, /t/=36, /k/=56 ms for French and /p/=61, /t/=75, /k/=79 ms for English. In the case of the English control group, the values were in the range of the reference values from the literature; however, the French control group measurements exceeded by more than 10 ms the monolingual reference values for French reported in the literature.

In order to compare the observed mean VOT durations for /p, t, k/ in the participants' L2 English and L3 French to the respective control groups, one-sample t-tests and a non-parametric Wilcoxon signed-rank matched pairs test were performed (assumed α=0.05).

Table 49. Comparison of mean VOT for /p, t, k/ in L2 English and L3 French to control groups values

Language	Parameter	VOT		
		/p/	/t/	/k/
English	English Controls	61	75	79
	L2 Mean	65	74.2	90.1
	SD	28.1	27.8	22
	p	0.111923	0.549199	0.000000*
French	French Controls	29	35	56
	L3 Mean	39.5	42.3	67.2
	SD	22.2	20.2	21.8
	p	0.000000*	0.000132*	0.000000*

*p<0.05

As can be seen in Table 49, in the case of L2 English the performed statistical tests did not point to any significant differences in the obtained measurements between the multilingual and the control group for /p/ and /t/, with overshoot values reported for /k/, i.e. 90 ms vs. 79 ms in the English controls. As stated earlier, the English controls' VOT values were well within the range reported in the literature, thus confirming the validity of the selected reference norms. As far as L3 French is concerned, statistically significant differences were found for the comparison of all the VOT durations, with the multilingual participants implementing the voiceless plosives with considerably longer lag values than the native French controls (i.e., 39.5 vs. 29 ms for /p/, 42 vs. 35 ms for /t/, 67 vs. 56 ms for /k/). On the whole, we can assume that the participants' L2 VOT values approximated those of the native English controls; however, there was a visible discrepancy in the case of L3 French values which deviated from the French control group.

5.4.1.7. VOT goodness of fit

In order to evaluate the participants' approximation to the target control VOT values, a measure for the nativeness effect was proposed. To this end, threshold levels were created at 10% intervals to reflect the degree of approximation or deviation from the control VOT values. Level 0 corresponded to a -10%÷10% deviation from the control mean VOT values, level 1 reflected a 11%÷20% deviation, level 2 a 21%÷30%, whereas level -1 corresponded to a -20%÷-11% deviation from the control values, etc. The participants' measurements were assigned to particular categories; the higher the number, the more they differed from the control baseline durations, with '+ values' indicating a longer lag and '- values' shorter VOT durations. Table 50 shows the distribution of percentage scores across the approximation levels for L2 English and L3 French in comparison to the control VOT durations.

Table 50. Level of approximation of L2 and L3 VOT values to control VOT (in % of participants)

Approximation level	L2 English fit to English control VOT			L3 French fit to French control VOT		
	/p/	/t/	/k/	/p/	/t/	/k/
-9	0	0	0	0	0	0
-8	0.4	0	0	0	0	0
-7	0.9	0.4	0	0.44	0	0
-6	4.4	3.1	0	3.08	0.9	0.5
-5	5.7	4.8	0.4	4.41	4.4	1.4
-4	7.0	10.6	1.8	7.49	4.4	1.9
-3	6.6	6.6	3.5	7.93	10.1	6.7
-2	8.3	10.1	3.1	5.29	5.7	5.3
-1	6.6	6.2	12.3	5.73	9.7	10.1
0	16.2	18.9	26.3	8.37	19.8	15.4
1	7.0	9.7	11.0	6.61	3.5	9.6
2	5.7	7.5	11.8	4.85	4.0	6.3
3	6.6	5.7	10.5	5.73	4.4	14.4
4	6.6	5.7	9.2	3.52	6.2	7.7
5	3.9	4.0	6.1	3.96	4.4	5.3
6	4.4	4.4	2.2	3.52	3.5	3.8
7	2.2	1.8	0.4	7.05	1.8	3.4
8	2.6	0.4	0.4	2.20	3.1	4.8
9	2.2	0	0.4	2.20	1.8	1.0
10	1.8	0	0.4	1.32	3.5	1.0
11	0.4	0	0	3.08	1.3	0.5
12	0.4	0	0	2.20	1.8	1.0
13	0	0	0	3.08	0.4	0

14	0	0	0	0.44	1.8	0
15	0	0	0	0.88	1.3	0
16	0	0	0	0.44	0.4	0
17	0	0	0	1.32	0.4	0
18	0	0	0	0.44	0.4	0
19	0	0	0	0.88	0.4	0
20	0	0	0	0.44	0.4	0
21	0	0	0	0.88	0	0
22	0	0	0	0.88	0	0
23	0	0	0	0.88	0	0
24	0	0	0	0.44	0	0

Table 51 presents the mean results of the approximation level analysis (i.e. the goodness of fit to the control VOT values) and points to diverse scores for L2 English (M=0.6) and L3 French (M=2.5). It appears that the nativeness effect based on the VOT approximation to the respective control values was more visible in the case of L2 English, in which the VOT measurements were within a close range of the control native values, whereas in the case of L3 French, the VOT durations deviated more significantly from the baseline values.

Table 51. Approximation level mean results

Language	N	Mean	SD	Median	Minimum	Maximum
L2	683	0.6	3.4	0.0	-8	14
L3	662	2.5	5.7	1.0	-7	27

5.4.1.8. Individual variation

The VOT measurements were also analyzed with respect to individual variation in the generated VOT values for /p/ /t/ /k/. The following figures illustrate the individual variation separately for the L1, L2 and L3 against the selected reference VOT values for particular languages.

As can be seen in Figures 28-30, the participants followed the universal VOT patterns in all their respective languages (L1, L2, L3), with bilabial plosives yielding the shortest VOT values, and velar – the longest.

The greatest variability in the VOT distribution can be observed in L3 French and L2 English. On the whole, the L3 values are overshoots of the reference French VOT durations, with the most extreme departures from the norm represented by such individuals as 1F_PS, 2F_PP, 2F_MBA, 2F_SF. Only in a few cases does the L3 performance approximate closely the native French VOT durations (i.e. 2F_ASN, 2F_IK, 2F_AJ). There was quite considerable variability exhibited also in L2 English, and the closest correspondence to the target English VOT durations was noticeable for such participants as 1F_BF, 1F_BK, 3F_AW. As expected, the VOT durations for L1 Polish exhibited less interspeaker variation and a fairly close correspondence to the VOT reference values for Polish.

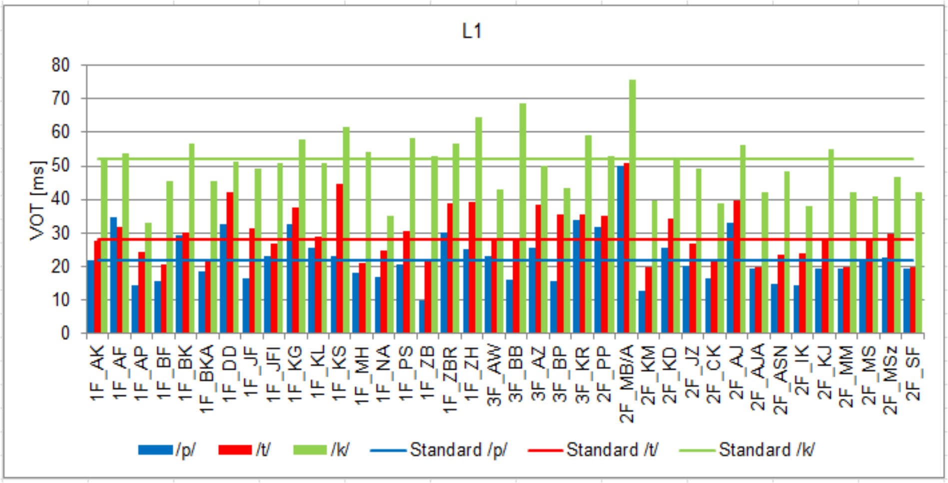


Figure 28. Individual variation in L3 French VOT for /p/ /t/ /k/ against the reference values

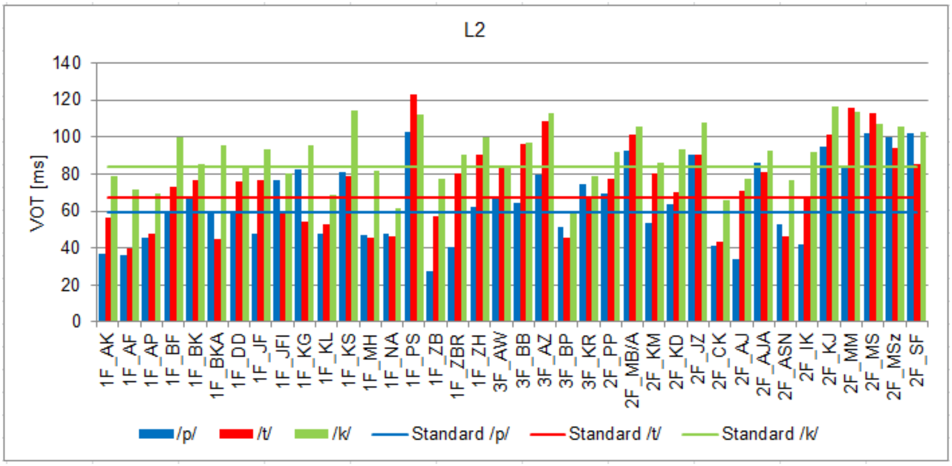


Figure 29. Individual variation in L2 English VOT for /p/ /t/ /k/ against the reference values

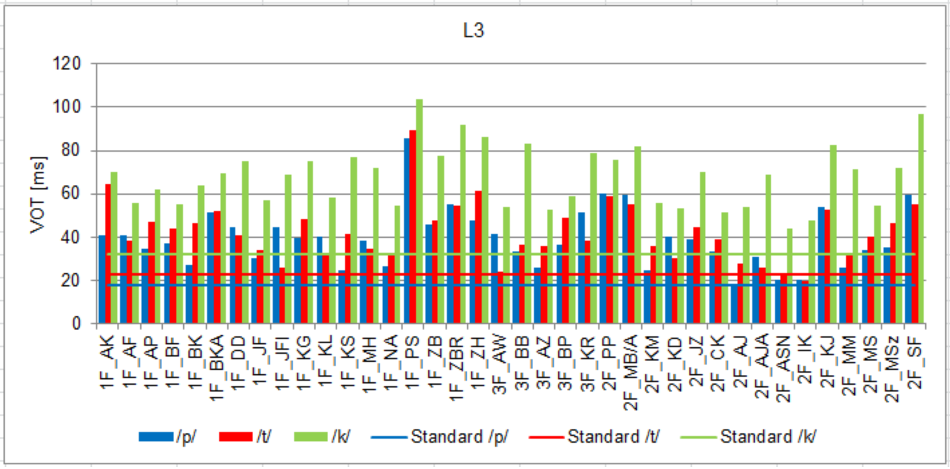


Figure 30. Individual variation in L1 Polish VOT for /p/ /t/ /k/ against the reference VOT values

5.4.1.9. Vocalic context effects

Non-parametric Kruskal-Wallis tests were used to investigate differences in the mean VOT with respect to the vocalic context (i.e. high, mid and low vowels) for the voiceless plosives under investigation, in each of the participants' languages. The assumed alpha level was α=0.05.

For the bilabial voiceless plosive /p/, the results of the performed Kruskal-Wallis Test point to statistically significant differences in the mean VOT values between /pa-, po/e-, pi-/ contrasts for all the three language systems. In L1 Polish, the low vowel /a/ and the mid vowels /o, e/ preceding the initial plosives, generated significantly shorter VOT durations compared to the high vowel /i/ context. In L2 English the only contrast that was statistically significant was that between the low and mid vowels contexts, with the former being longer in duration. The pattern observed in L3 French was similar to that of L1 Polish, with the high vowel context yielding a significantly longer VOT than the low or mid vowel contexts (see Table 52) .

Table 52. Comparison of mean VOT between /pa-, po/e-, pi-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	pa- vs po/e-	pa- vs pi-	po/e- vs pi-
L1	0.0000*	1.000000	0.000013*	0.000633*
L2	0.0012*	0.000714*	0.232915	0.168295
L3	0.0003*	1.000000	0.001350*	0.001285*

*p<0.05

The analysis of the vocalic context effects for the alveolar plosive /t/ demonstrated statistically significant differences in the mean VOT values for all the languages, i.e. L1 Polish, L2 English, and L3 French. The observed pattern was the same in all the language systems, with the low and mid vowel contexts generating shorter VOT durations compared to the high vowel /i/ contexts, yet there was no significant difference in the VOT measures between the low and mid vowels contexts for /t/ (see Table 53).

Table 53. Comparison of mean VOT between /ta-, to/e-, ti-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ta- vs to/e-	ta- vs ti-	to/e- vs ti-
L1	0.0000*	0.695276	0.000000*	0.000000*
L2	0.0073*	1.000000	0.014175*	0.028646*
L3	0.0000*	1.000000	0.000001*	0.000000*

*p<0.05

The results of the Kruskal-Wallis test for the multiple comparisons between the vocalic contexts for the velar plosive /k/ showed statistically significant differences in the mean VOT measures for all the languages involved. For L1 Polish and L3 French, similarly to the previous results, the high /i/ vowel context resulted in longer VOT durations of the preceding consonant compared to both the low and mid vowel contexts. A slightly different pattern was observed in the case of L2 English, where significant differences in the VOT durations were found for the comparisons between the high and mid vowels as well as the low and mid vowels (see Table 54).

Table 54. Comparison of mean VOT between /ka-, ko/e-, ki-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ka- vs ko/e-	ka- vs ki-	ko/e- vs ki-
L1	0.0000*	1.000000	0.000000*	0.000000*
L2	0.0000*	0.000230*	0.123501	0.000000*
L3	0.0000*	0.842127	0.000000*	0.000067*

*p<0.05

All in all, the vocalic context effects observed in the data follow, to a large extent, the universal trends in the VOT durations, where the context of high vowels generates longer VOT values in the preceding plosives than the context of low vowels. The exhibited patterns were especially consistent in L1 Polish and L3 French, while L2 English demonstrated fewer significant differences in the vocalic effects (see Figures 31-33 presenting box plots of the mean VOT values for various vocalic contrasts in the three respective languages).

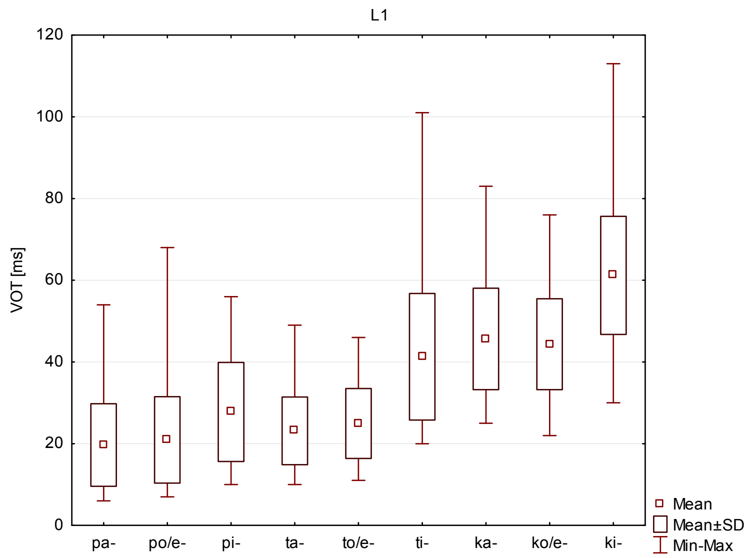


Figure 31. Box plot of mean VOT values for various vocalic contrasts in L1 Polish

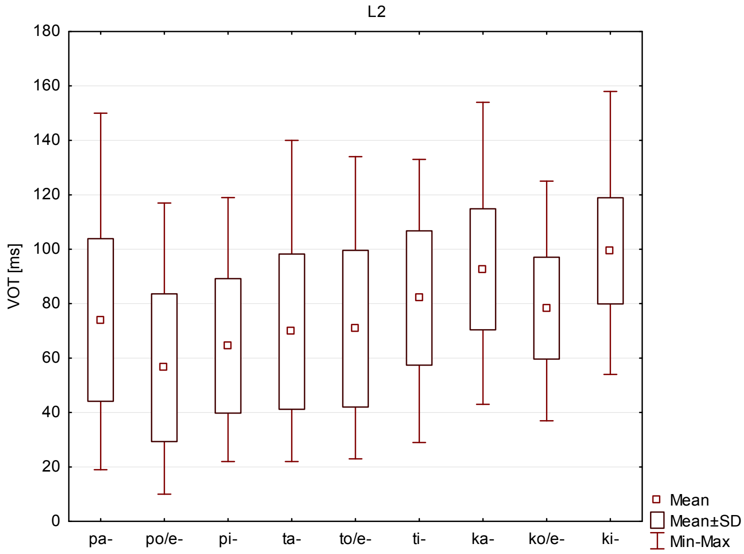


Figure 32. Box plot of mean VOT values for various vocalic contrasts in L2 English

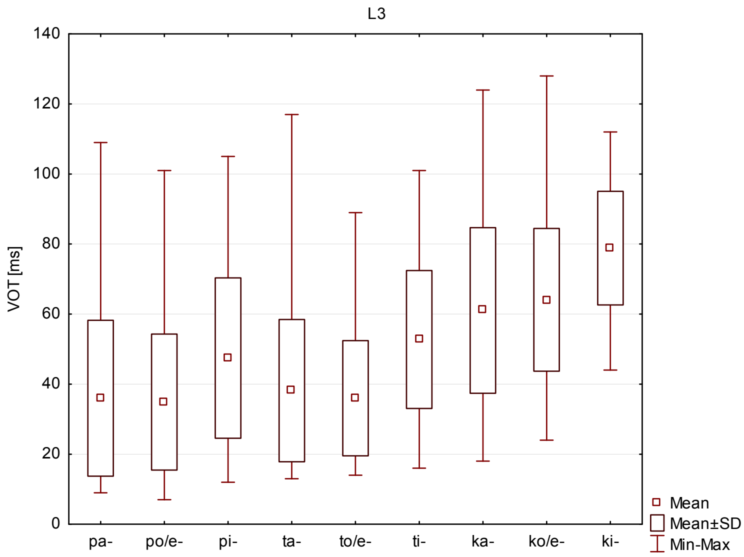


Figure 33. Box plot of mean VOT values for various vocalic contrasts in L3 French

5.4.1.10. Analysis of variance

The performed analysis of variance included a two-factor ANOVA between languages (L1, L2, L3) and the VOT durations of the voiceless plosive sounds /p, t, k/ (see Table 55).

Table 55. Two-factor analysis of variance for /p, t, k/ and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	704.34 (0.05;2;2020)	0.000000*
plosive (/p/, /t/, /k/)	298.3 (0.05;2;2020)	0.000000*
Language*plosive	2.77 (0.05;4;2020)	0.026071*

*p<0.05

The results show that the differences in the VOT values within the factors of language (F (2; 2020)=704 p<.05) and plosives (F (2; 2020)=298, p<.05) were found to be significant. Moreover, the interaction between the languages and the plosives on the VOT values was also shown to be significant (F (4; 2020)=2.77, p<.05). The findings indicate that there are significant differences between the VOT values for at least two languages, and for at least two plosives. The existing interaction between the languages and the plosives depends on the type of language (L1, L2, L3), as presented in Figure 34.

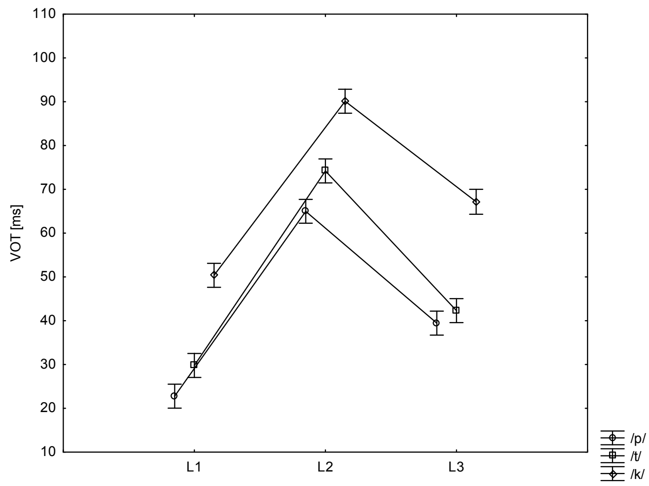


Figure 34. Interaction between the language and plosive factors

Another analysis investigated the interaction of the ensuing vowel context and the language on the observed VOT durations. To this end, a two-factor analysis ANOVA was performed for the factors of the language (L1, L2, L3) and the context of the vowel (i.e. low _/a/, high _/i/, and mid _/e, o/) following the voiceless plosives in the stressed onset positions in the target words. The results demonstrate that there are significant differences in the VOT values within the factor of the languages (F (2; 2020)=581, p<.05) and the vowel context (F (2; 2020)=60.8, p<.05). It can be interpreted that significant differences between the VOT values were found for at least two languages, and for at least two vowel contexts. Moreover, a significant interaction was found between the two factors (F (4; 2020)=5, p<.05), which depends on the type of the language (see Table 56 and Figure 35).

Table 56. Two-factor analysis of variance for vowel context /-a-/, /-o/e-/, /-i-/ and L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	581.12 (0.05;2;2020)	0.000000*
Vowel (/-a-/. /-o/e-/. /-i-/)	60.83 (0.05;2;2020)	0.000000*
Language*Vowel	5.06 (0.05;4;2020)	0.000466*

*p<0.05

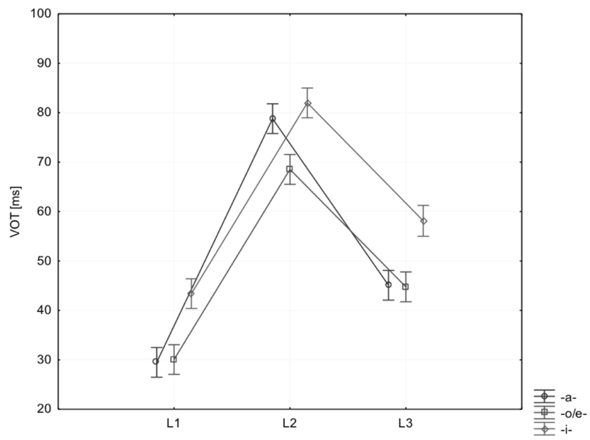


Figure 35. Interaction between the language and vowel context factors

The observed patterns of interaction were generally consistent with the universal effects of the vocalic context on the duration of the preceding plosive. The high vowel context generated the longest VOT values in all the languages. The differences in durations between the mid and low vowel context was negligible in the case of L1 Polish and L3 French, whereas in L2 English the pattern was slightly different to the expected one, with the low vowel context generating a longer VOT than the mid vowel context.

Another analysis was performed to investigate the language proficiency level and the language (L1, L2, L3). To this end, the participants were subdivided into two proficiency subgroups 1F vs. 2/3F, the former representing a lower proficiency level in L3 French than the latter subgroup. The results of the two-factor ANOVA for the proficiency groups and for L1, L2, L3 demonstrated that the differences in the VOT values within the factors of the language ($F(2; 2023)=547.8$, $p<.05$) and with respect to the proficiency group ($F(2; 2023)=4.35$, $p<.05$) were found to be significant. Moreover, there was a significant interaction between the factors of the language and the group on the mean VOT values ($F(2; 2023)=26.6$, $p<.05$) (see Table 57 and Figure 36).

Table 57. Two-factor analysis of variance for proficiency groups and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	547.82 (0.05;2;2023)	0.000000*
Group (1F vs. 2/3F)	4.35 (0.05;1;2023)	0.037042*
Language*Group	26.60 (0.05;2;2023)	0.000000*

*p<0.05

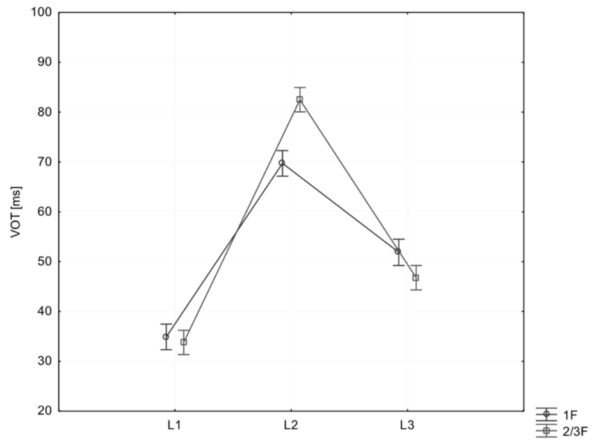


Figure 36. Interaction between the language and proficiency group factors

A subsequent two-factor analysis of variance ANOVA for the proficiency subgroups and the place of articulation PoA /p, t, k/ performed separately for all the languages (L1, L2, L3) showed the following patterns. For L1 Polish, there were statistically significant differences between the mean VOT values for at least two places of articulation (F (2; 678) =259, p<.05); there were no significant differences in VOT for the proficiency groups; there were no significant interactions between PoA and proficiency groups.

For L2 English and L3 French, there were statistically significant differences between the mean VOT values for at least two places of articulation, i.e. for L2 (F (2; 677)=58.51, p<.05) and for L3 (F (2; 656) =109.7, p<.05); there were statistically significant differences between the mean VOT values for the proficiency group, i.e. for L2 (F (2; 677)=43, p<.05) and for L3 (F (2; 656) =12, p<.05); there were no significant interactions between PoA and groups (see Table 58 and Figure 37).

Table 58. Two-factor analysis of variance for proficiency groups and place of articulation PoA /p, t, k/ and for languages L1, L2, L3

Factor	L1		L2		L3	
	F	p	F	p	F	p
PoA (/p/, /t/, /k/)	258.99 (0.05;2;678)	0.000000*	58.54 (0.05;2;677)	0.000000*	109.71 (0.05;2;656)	0.000000*
Group (1F, 2/3F)	1.12 (0.05;1;678)	0.290164	43.37 (0.05;1;677)	0.000000*	12.16 (0.05;1;656)	0.000521*
PoA*Group	0.51 (0.05;2;678)	0.601551	2.43 (0.05;2;677)	0.089202	0.40 (0.05;2;656)	0.669627

*p<0.05

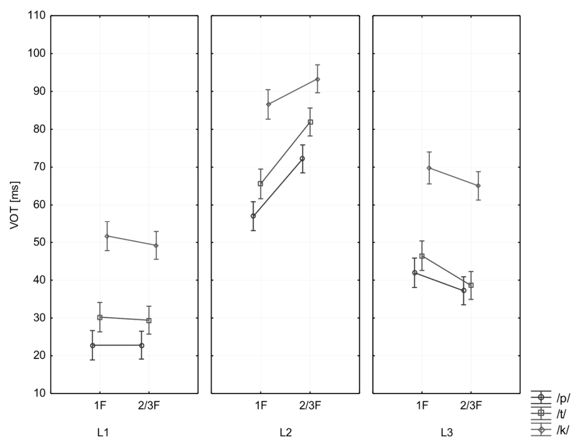


Figure 37. Interaction of PoA (/p/, /t/, /k/) and Group (1F, 2/3F) for L1 Polish, L2 English and L3 French

5.4.1.11. Multiple regression analysis

Multiple regression analyses were performed to estimate the relationship between the dependent and independent variables in Group A.

The first regression analysis investigated the relationship between the VOT values as the dependent variable and a number of independent variables including the language (L1, L2, L3), the place of articulation PoA (/p/, /t/, /k/), the vowel context and the proficiency group. The assumed levels of significance point to a statistically significant interdependence between the VOT values and the following variables of the language, the place of articulation, the vowel context and the group. The R squared result indicated that 53.7% of the variance is accounted for by the independent variables.

The beta coefficient was calculated to establish which of the independent variables had a greater effect on the dependent variable in the multiple regression analysis. The beta coefficient values suggest that "Language 2", i.e. L2 (b=0.6) and "Language 1" (L1, L3) (b=-0.5) as well as PoA 1, i.e. /p/ (b=-0.3) have the biggest influence on the VOT values, whereas the impact of the remaining variables is relatively small (see Tables 59 and 60).

Table 59. Multiple regression analysis 1

Variable	R	R^2	F	p
VOT	0.73	0.537	335.11 (0.05;7;2021)	0.000000*

*p<0.05

Table 60. Results of the multiple regression 1 for variables

Effect	t	p	Beta
Target word	118.94	0.000000*	–
Language 1	-30.32	0.000000*	-0.527
Language 2	36.25	0.000000*	0.630
Group	-2.22	0.026305*	-0.034
PoA 1	-17.54	0.000000*	-0.305
PoA 2	-7.49	0.000000*	-0.130
Vowel 1	-3.76	0.000174*	-0.066
Vowel 2	-8.95	0.000000*	-0.156

*p<0.05

A further multiple regression was performed for the L3 VOT values as the dependent variable, and L1 VOT and L2 VOT as independent variables. The analysis pointed to a statistically significant interdependence among the variables under investigation. The regression model is explained in 31.9% by the independent variables L1 VOT and L2 VOT (R^2=0.319). The bigger impact on the dependent variable L3 VOT is exerted by the L1 VOT values (b=0.47) than by the L2 VOT (b=0.19) (see Tables 61 and 62).

Table 61. Multiple regression analysis 2

Variable	R	R^2	F	p
L3 VOT	0.56	0.319	154.09 (0.05;2;658)	0.000000*

*p<0.05

Table 62. Results of the multiple regression 2 for variables

Effect	t	p	Beta
Target word	5.73	0.000000*	–
L1 VOT	13.62	0.000000*	0.469
L2 VOT	5.54	0.000000*	0.191

*p<0.05

Finally, a regression analysis was calculated for the L3 VOT values as the dependent variable and other independent variables, including L1 VOT, L2 VOT, the place of articulation /p, t, k/, the vowel context and the proficiency group. The results show that the regression model is explained in 40.5% of cases by the independent variables. The assumed level of testing probability points to a statistically significant impact of such independent variables as L1 and L2 VOT, the place of articulation /p, t, k/ and Vowel 1, i.e. /-a, -i/. The greatest impact on the L3 VOT is exerted by L1 VOT (b=0.22) and L2 VOT (b=0.21), whereas the remaining variables have a smaller influence (see Tables 63 and 64).

Table 63. Multiple regression analysis 3

Factor	R	R^2	F	p
L3 VOT	0.64	0.405	63.55 (0.05;7;653)	0.000000*

*p<0.05

Table 64. Results of multiple regression for variables

Effect	t	p	Beta
Target word	8.35	0.000000*	–
L1 VOT	4.84	0.000002*	0.217
L2 VOT	6.22	0.000000*	0.215
Group	5.48	0.000000*	0.172
PoA 1	-3.84	0.000135*	-0.162
PoA 2	-5.46	0.000000*	-0.194
Vowel 1	-3.66	0.000272*	-0.132
Vowel 2	-1.94	0.053296	-0.070

*p<0.05

5.4.2. Results for Group B

In the following sections, the results for Group B with L1 Polish, L2 English and L3 German, will be presented in detailed.

5.4.2.1. Mean VOT values for L1, L2 and L3

First, the Shapiro–Wilk test of normality was administered to verify the null hypothesis for this test that the VOT measurements data are normally distributed. Since the chosen alpha level was 0.05 and the p-values for the VOT variable in several categories were less than 0.05, then the null hypothesis that the data are normally distributed had to be rejected. Consequently, non-parametric tests were applied for the subsequent statistical analyses of the data in Group B.

Tables 65-67 present the mean results of the VOT measurements of the voiceless plosives /p/, /t/, /k/ of the target words read in the carrier phrasesin the participants' L1 Polish, L2 English and L3 German.

Table 65. Mean VOT values for target words in L1 Polish

Words in L1	N	Mean	SD	Median
pan	26	29.2	16.1	23.0
para	26	24.4	12.1	23.0
pech	26	21.0	9.8	17.0
pole	26	29.5	12.8	25.0
pik	26	28.3	13.3	26.5

piwo	26	36.7	12.5	33.0
tak	26	23.7	8.6	22.5
tango	26	28.3	7.5	24.5
tor	26	31.9	10.7	32.5
testy	26	26.3	7.1	24.5
tik	26	37.8	12.5	36.0
tiry	26	49.4	14.7	45.0
kat	26	52.1	11.4	55.0
kanon	26	53.8	12.2	52.5
kot	26	45.7	11.1	43.5
kelner	26	48.8	10.0	49.5
kit	26	71.0	13.6	67.0
kino	26	64.4	8.6	67.0

Table 66. Mean VOT values for target words in L2 English

Words in L2	N	Mean	SD	Median
pan	26	71.5	32.4	64.5
party	26	55.5	27.9	55.0
pot	26	63.9	31.7	59.0
pencil	26	43.5	26.0	43.0
peace	26	62.1	30.2	60.5
Peter	26	48.8	27.2	39.5
task	26	72.0	28.0	71.0
taxi	26	65.3	30.5	69.0
test	26	70.0	24.8	65.5
toffee	26	67.4	25.6	70.0
team	25	93.2	29.9	89.0
teacher	26	76.2	27.1	76.0
cat	26	98.5	26.9	101.5
carpark	26	92.5	24.8	97.0
kept	26	77.0	22.4	81.0
coffee	26	76.4	23.6	69.0
keen	26	98.3	25.7	92.0
keeper	25	92.4	20.2	93.0

Table 67. Mean VOT values for target words in L3 German

Words in L3	N	Mean	SD	Median
Paar	26	52.0	22.6	52.5
Panne	26	46.6	23.9	48.0
Pelz	26	38.1	19.0	36.0

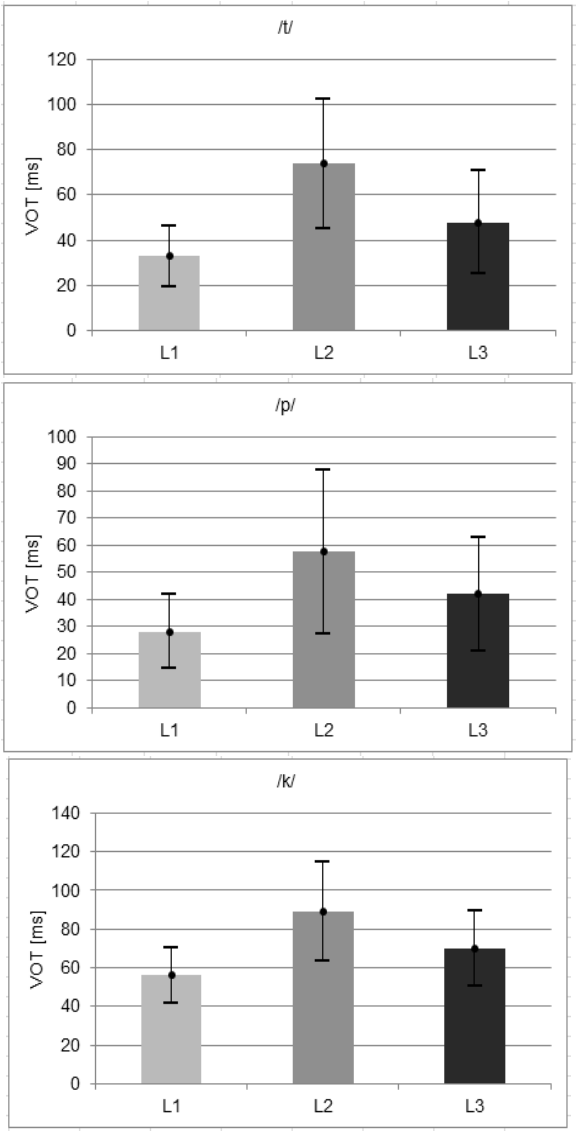
Pose	26	44.8	19.2	46.0
Pirsch	26	38.8	20.6	33.0
Pinsel	26	31.3	12.5	30.5
Tanz	26	50.5	28.5	48.0
Tasse	26	50.1	24.1	52.5
Text	26	34.8	13.8	37.5
Tochter	24	39.8	19.6	37.5
tief	23	68.1	19.3	69.0
ticken	26	45.7	14.3	43.0
Kalk	25	74.8	27.1	71.0
Karte	26	66.5	21.6	64.5
kommen	26	66.0	12.2	67.5
Kegel	25	60.8	12.2	61.0
Kind	26	64.8	15.4	62.0
Kilo	26	86.8	11.4	86.5

The results of the acoustic measurements of the mean voice onset time for the voiceless plosives /p/, /t/, /k/ in the stress onset positions in the participants' L1 Polish, L2 English and L3 German are presented in Table 68 and Figures 38-40.

The conducted durational measurements demonstrate that the multilingual participants produced the voiceless plosives in the stress onset positions with the mean voice onset time values that were the shortest in L1 Polish (/p/=28 ms, /t/=33 ms, /k/=56 ms), the longest in L2 English (/p/=57.5 ms, /t/=74 ms, /k/=89 ms) and with in between values for L3 German (/p/=42 ms, /t/=48 ms, /k/=70 ms).

Table 68. Mean VOT values for L1, L2 and L3

VOT	Language	N	Mean	SD	Median
/p/	L1	156	28.2	13.6	25.0
	L2	156	57.5	30.4	54.5
	L3	156	41.9	20.8	38.5
/t/	L1	156	32.9	13.5	30.0
	L2	155	73.9	28.8	73.0
	L3	151	47.9	22.7	43.0
/k/	L1	156	56.0	14.2	58.0
	L2	155	89.2	25.4	87.0
	L3	154	70.0	19.4	67.0



Figures 38-40. VOT measurements for /p, t, k/ in L1 Polish, L2 English and L3 German

5.4.2.2. Cross-language comparison of VOT means

The first series of statistical tests was conducted in order to investigate the language effect in the VOT measurements in Group B. Due to the lack of normal distribution, a non-parametric Kruskal-Wallis ANOVA was used to compare the mean VOT values for /p/ /t/ /k/ across the languages of the multilingual participants. The assumed alpha level was 0.05.

The performed ANOVA pointed to significantly different values for all the initial voiceless plosives between all the language pairings involved, i.e. L1 Polish vs. L2 English, L1 Polish vs. L3 German as well as L2 English vs. L3 German ($p<.05$), (see Table 69).

Table 69. Mean VOT comparison for /p t k/ in L1 Polish, L2 English, L3 German

Kruskal-Wallis Test	p for multiple comparisons		
	L1 vs L2	L1 vs L3	L2 vs L3
/p/	0.000000*	0.000000*	0.000065*
/t/	0.000000*	0.000000*	0.000000*
/k/	0.000000*	0.000000*	0.000000*

*$p<0.05$

In the individual analyses of measurements for /p/, /t/, /k/, the p values for the multiple comparisons point to statistically significant differences in the mean VOT between all the languages involved, i.e. betweenL1 and L2 (mean VOT values in L1 Polish are lower than in L2 English); between L1 and L3 (mean VOT values in L1 Polish are lower than in L3 German); between L2 and L3 (mean VOT values in L2 English are higher than in L3 German).

Summing up, the performed tests yielded cross-linguistic differences that proved to be significant between the native tongue (Polish) and both non-native languages (L2 English and L3 German). Moreover, similar cross-linguistic patterns were observed in the pairwise comparison of means between L2 English and L3 German. In conclusion, the language effect was thus observed to hold between all the respective language systems.

The following box plots (Figures 41-43) illustrate the observed tendencies in the VOT patterns in the respective languages separately for the stressed onset plosives /p/ /t/ and /k/. It is interesting to observe that VOT in L3 German demonstrates in-between values between L1 Polish and L2 English with respect to the VOT means, the standard deviation as well as the minimum-maximum range. In L1 Polish it is not only the mean VOT that remains significantly lower, but it is also the standard deviation and the minimum-maximum range that is smaller, compared especially against L2 English, in which all these parameters are more extreme.

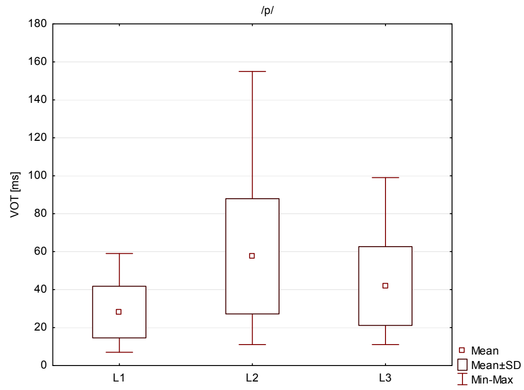


Figure 41. Box plot of mean VOT values for /p/ in L1 Polish, L2 English and L3 German

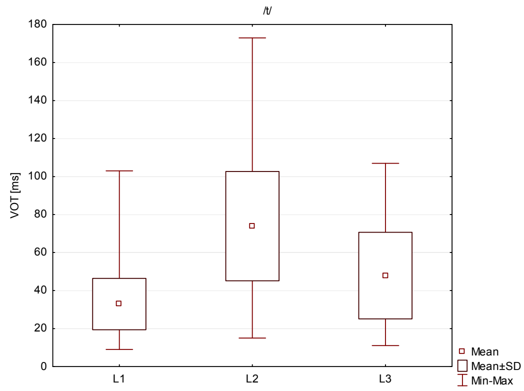


Figure 42. Box plot of mean VOT values for /t/ in L1 Polish, L2 English and L3 German

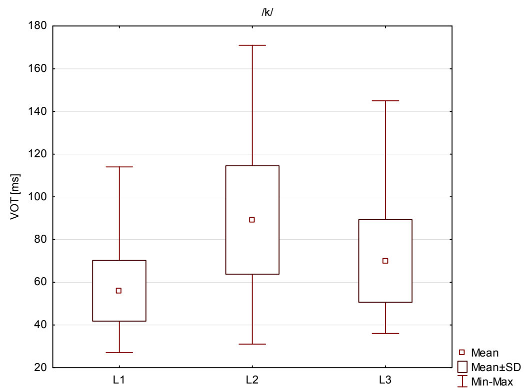


Figure 43. Box plot of mean VOT values for /k/ in L1 Polish, L2 English and L3 German

5.4.2.3. Proficiency group effect

In order to examine the language proficiency group effect on the VOT acquisition patterns in Group B, the non-parametric Mann-Whitney U tests were performed for the mean VOT values for /p/, /t/, /k/ in L1 Polish, L2 English and L3 German for the two proficiency subgroups, i.e. lower proficiency 1/2G (N=16) vs. higher proficiency 3G (N=10) (see Table 70).

The results of the analysis showed only one statistically significant difference between the VOT values with respect to the language proficiency group for /k/ in L2 English, i.e. the mean VOT in the 1/2G group is lower than in the 3G group. In all the remaining cases, no significant differences were reported between the results of the two L3 proficiency subgroups in Group B.

Table 70. Results of Mann-Whitney test between two proficiency groups in L3 German for /p, t, k/ for L1, L2 and L3

Sounds	Language	Z corr.	p
/p/	L1	-0,5	0,639868
	L2	1,5	0,125421
	L3	-0,7	0,463834
/t/	L1	-0,1	0,947489
	L2	1,8	0,066413
	L3	-1,0	0,323530
/k/	L1	-0,8	0,415316
	L2	2,8	0,004739*
	L3	-0,8	0,435668

*$p<0.05$

Recapitulating, as far as the proficiency group effect is concerned, no regularities were observed to hold with respect to the mean VOT values for /p/, /t/, /k/ in all the languages concerned. It follows that the proficiency level in L3 German was not a significant determiner of the respective VOT durations in Group B.

The following box plots (Figures 44-46) illustrate the observed tendencies in the group comparison based on the L3 proficiency criteria (A- 1/2G – less advanced, B- 3G – more advanced group).

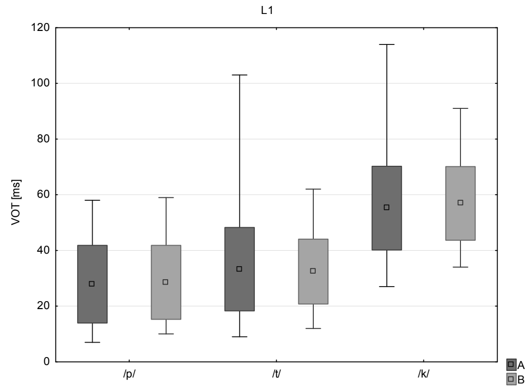


Figure 44. Box plot of mean VOT values for /p/, /t/, /k/ in L1 Polish – proficiency group comparison

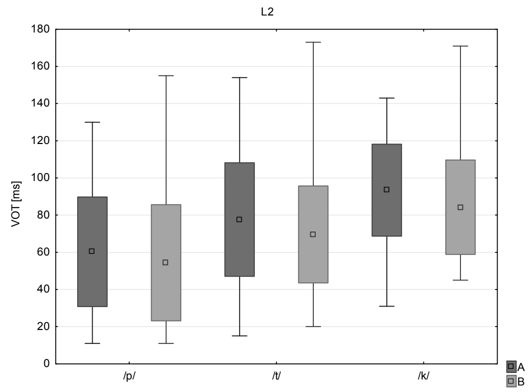


Figure 45. Box plot of mean VOT values for /p/, /t/, /k/ in L2 English – proficiency group comparison

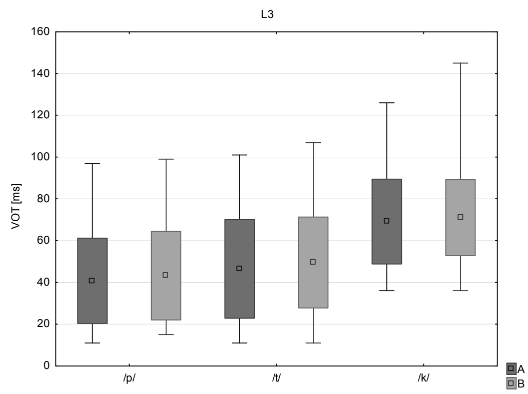


Figure 46. Box plot of mean VOT values for /p/, /t/, /k/ in L3 German – proficiency group comparison

5.4.2.4. Cross-linguistic correlations between VOT values

In order to investigate the dependence between two variables, i.e. the observed VOT values in L3 German and those of the native L1 Polish as well as between L3 German and L2 English, the Pearson product-moment correlation analysis was applied (α=0.05). The calculated coefficients pointed to prevailingly positive weak to moderate correlations between the mean VOT values in Group B.

For the voiceless bilabial plosive /p/, statistically significant correlations were found for both pairs of variables. There was a weak positive correlation (R=0.23) between L1 Polish and L3 German, whereas the correlation in the VOT values for /p/ between the non-native languages (L2 English and L3 German) was moderate (R=0.5) (see Table 71).

Table 71. Pearson's correlation for /p/

Pair of variables for /p/	N	R	t	p
VOT L1 & VOT L3	156	0.23	2.94	0.003763*
VOT L2 & VOT L3	156	0.50	7.21	0.000000*

*p<0.05

As far as Pearson's correlations for /t/ are concerned, a strong positive correlation was found between L2 English and L3 German (R=0.59); however, there was no significant correlation for the second pair of variables, i.e. L1 Polish and L3 German (see Table 72). The correlation is illustrated in Figure 47 presenting an example of a scatterplot.

Table 72. Pearson's correlation for /t/

Pair of variables for /t/	N	R	t	p
VOT L1 & VOT L3	150	0.13	1.61	0.109278
VOT L2 & VOT L3	150	0.59	8.98	0.000000*

*p<0.05

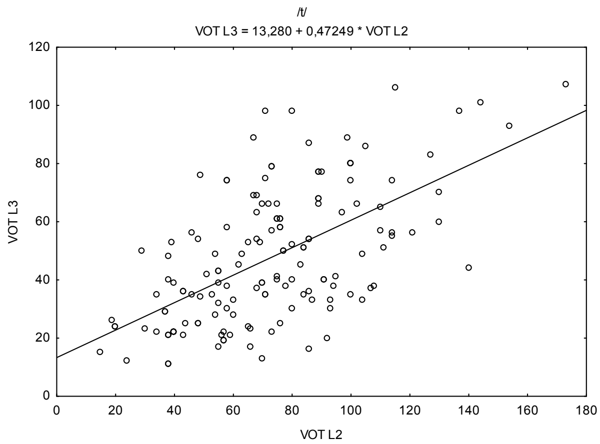


Figure 47. Scatterplot of VOT for /t/ between L2 English and L3 German

For the velar plosive /k/ the obtained Pearson's correlations for both pairs of variables were statistically significant; with the coefficient ranging from weak for L1 Polish vs. L3 German (R=0.29) to moderate for L2 English vs. L3 German (R=0.46) (see Table 73).

Table 73. Pearson's correlation for /k/

Pair of variables for /k/	N	R	t	p
VOT L1 & VOT L3	153	0.29	3.69	0.000310*
VOT L2 & VOT L3	153	0.46	6.41	0.000000*

*p<0.05

The results of the correlation analyses point out that the VOT values in L3 German were correlated more consistently and more strongly with those in L2 English, whereas the covariance between L3 German and L1 Polish remained much weaker in Group B.

5.4.2.5. Comparison to VOT reference values

In order to compare the acoustic measurements obtained in the study to the VOT reference values from the literature for the respective languages, a series of analyses was performed. One-sample t-tests and a non-parametric Wilcoxon matched pairs test were administered to compare the calculated mean VOT durations for /p, t, k/ in L1 Polish, L2 English and L3 German to the monolingual reference values, as quoted in the literature, i.e. Keating et al. 1981 for Polish, Lisker and Abramson 1964 for English, and Angelowa and Pompino-Marschall 1985 for German.

Table 74. Comparison to VOT reference values in Polish, English and German ([1]Keating et al., 1981; [2]Lisker and Abramson, 1964; [3]Angelowa and Pompino-Marschall, 1985)

Language	Parameter	VOT		
		/p/	/t/	/k/
Polish	Ref. VOT[1]	22	28	52
	L1 Mean	28.2	32.9	56.0
	SD	13.6	13.5	14.2
	p	0.000014*	0.000212*	0.001524*
English	Ref. VOT[2]	59	67	84
	L2 Mean	57.5	73.9	89.2
	SD	30.4	28.8	25.4
	p	0.141282	0.014729*	0.036078*
German	Ref. VOT[3]	36	39	47
	L3 Mean	41.9	47.9	70.0
	SD	20.8	22.7	19.4
	p	0.010273*	0.000110*	0.000000*

*$p<0.05$

The findings showed that nearly all the mean VOT measurements for L1 Polish and L2 English and L3 German differed significantly from the monolingual native norms from the literature (see Table 74).

Interestingly, in L1 Polish, the VOT durations of all the three voiceless stops were longer than the monolingual norms (cf. Keating et al., 1981), although still within the accepted 5-10 ms range (/p/ 28 vs. 22 ms, /t/ 33 vs. 28 ms, /k/ 56 vs. 52 ms). The overshoots of the native VOT values could have resulted from the influence of the long-lag values in the non-native languages, particularly the well established VOT values for L2 English.

As far as the VOT in L2 English is concerned, the reported measures for /p/ corresponded closely to the reference values, yet, those for /t/ and /k/ exhibited longer mean durations of voice onset time than those reported in the monolingual literature references (cf. Lisker and Abramson, 1964), i.e. /t/ 74 vs. 67 ms; /k/ 89 vs. 84 ms. Similarly to L1 Polish, the differences in the VOT durations did not exceed 5 ms. In turn, the measurements for L3German were found to depart the most from the monolingual VOT norms (cf. Angelowa and Pompino-Marschall, 1985). The bilabial, alveolar and velar voiceless plosives were realized, on average, with a longer lag than the reference values (/p/ 42 vs. 36 ms; /t/ 48 vs. 39 ms; /k/ 70 vs. 47 ms).

Summing up, nearly all of the observed VOT values of the multilingual participants in Group B departed from the reference norms and exhibited lengthening; however, in the case of L1 Polish and L2 English they could be regarded as overshoots approximating the target values within the acceptable 5 ms range. On the other hand, a considerable VOT lengthening was observed for L3 German when compared to the literature reference values. The German stops were implemented by the multilingual participants with a longer lag and thus the L3 phonetic norms were not approximated fully. An interesting observation is that the L1 VOT values seem to have been affected, to some extent, by the L2 long lag durations, thus providing evidence for the regressive transfer.

5.4.2.6. Comparison to Control groups

To verify further the reference values presented in the literature, further VOT measurements were generated from the control groups involved in the study. To this end, an identical recording procedure was followed as in the case of the experimental group with the exception that the controls performed only in their native tongue. For the purpose of the VOT comparisons in Group B, the German and English control groups were used. The German control group consisted of 17 native Germans (i.e. students and lecturers at the University of Leipzig), who were recorded reading the word list in the carrier phrases for German. The English control group included 17 native English speakers (i.e. lecturers at the University of Essex and Adam Mickiewicz University in Poznań), who were recorded reading the list for English.

The generated control mean values were /p/=62, /t/=67, /k/=82 ms for German, and /p/=61, /t/=75, /k/=79 ms for English. In the case of the English control group, the VOT values were in the range of the reference values from the literature; however, the German control group measurements exceeded considerably the reference values for German from the literature.

One-sample t-tests and a non-parametric Wilcoxon matched pairs test were performed to compare the observed mean VOT durations for /p, t, k/ in the participants' L2 English and L3 German to the respective control groups (assumed α=0.05). Statistically significant differences were found for the comparisons of the VOT durations of nearly all the plosives with the control native groups. The L2 English values were slightly shorter than the English native controls for /p/ and /t/, but longer for /k/. In the case of L3 German the observed values were considerably shorter when compared to those of the German native controls. The multilingual participants did not approximate fully the native VOT targets in their L3 German, instead they implemented them as compromise values intermediate between their L1 and the German target norms (see Table 75).

Table 75. Comparison of mean VOT for /p, t, k/ in L2 English and L3 German to control groups values

Language	Parameter	VOT		
		/p/	/t/	/k/
English	English Controls	61	75	79
	L2 Mean	57.5	73.9	89.2
	SD	30.4	28.8	25.4
	p	0.027142*	0.320026	0.000014*
German	German Controls	62	67	82
	L3 Mean	41.9	47.9	70.0
	SD	20.8	22.7	19.4
	p	0.000000*	0.000000*	0.000000*

*p<0.05

5.4.2.7. VOT goodness of fit

In order to evaluate the participants' approximation to the target control values in Group B, the same procedure was followed as in Group A. The calculation of the measure of the nativeness effect or VOT goodness of fit, based on the degree of approximation or deviation from the control VOT values, is presented in section 2.1.7. Table 76 shows the distribution of percentage scores across the approximation threshold levels for L2 English and L3 German in comparison to the respective control VOT durations.

Table 76. Level of approximation of L2 and L3 VOT values to control VOT (in % of participants)

Approximation level	L2 English to Control VOT			L3 German to Control VOT		
	/p/	/t/	/k/	/p/	/t/	/k/
-8	1.9	0	0	1.9	2.6	0
-7	5.1	2.6	0	7.7	4.6	0
-6	5.8	1.3	0.6	17.9	14.6	0
-5	7.1	3.2	0.6	8.3	7.9	2.6
-4	8.3	10.3	1.3	10.3	17.2	9.1
-3	9.6	5.2	4.5	10.9	5.3	16.9
-2	6.4	11.6	7.7	10.9	9.3	17.5
-1	5.8	6.5	12.9	6.4	9.9	17.5
0	14.1	23.2	21.9	13.5	11.3	22.1
1	7.7	10.3	11.6	5.8	9.3	7.8
2	10.9	5.8	9.0	2.6	2.0	1.9
3	1.9	7.1	12.9	0	2.0	1.3
4	1.3	3.9	3.9	1.3	2.0	1.3
5	3.8	2.6	4.5	2.6	2.0	1.3
6	2.6	1.3	3.2	0	0	0
7	0.6	1.9	1.9	0	0	0.6
8	1.9	1.3	1.9	0	0	0
9	0.6	0.6	0.6	0	0	0
10	1.9	0.6	0.6	0	0	0
11	0.6	0.6	0	0	0	0
12	0.6	0	0	0	0	0
13	0.6	0	0	0	0	0
14	0.6	0	0	0	0	0

Summing up, Table 77 presents the mean results of the approximation level analysis for Group B points to different patterns of approximation to the respective VOT control values. For L2 English, a strong L1 target nativeness effect was exhibited with on average 2% departure from the control VOT, whereas in the case of L3 German the difference was more outstanding, i.e. the VOT durations were approximately 22% shorter than the German control baseline values.

Table 77. Approximation level mean results

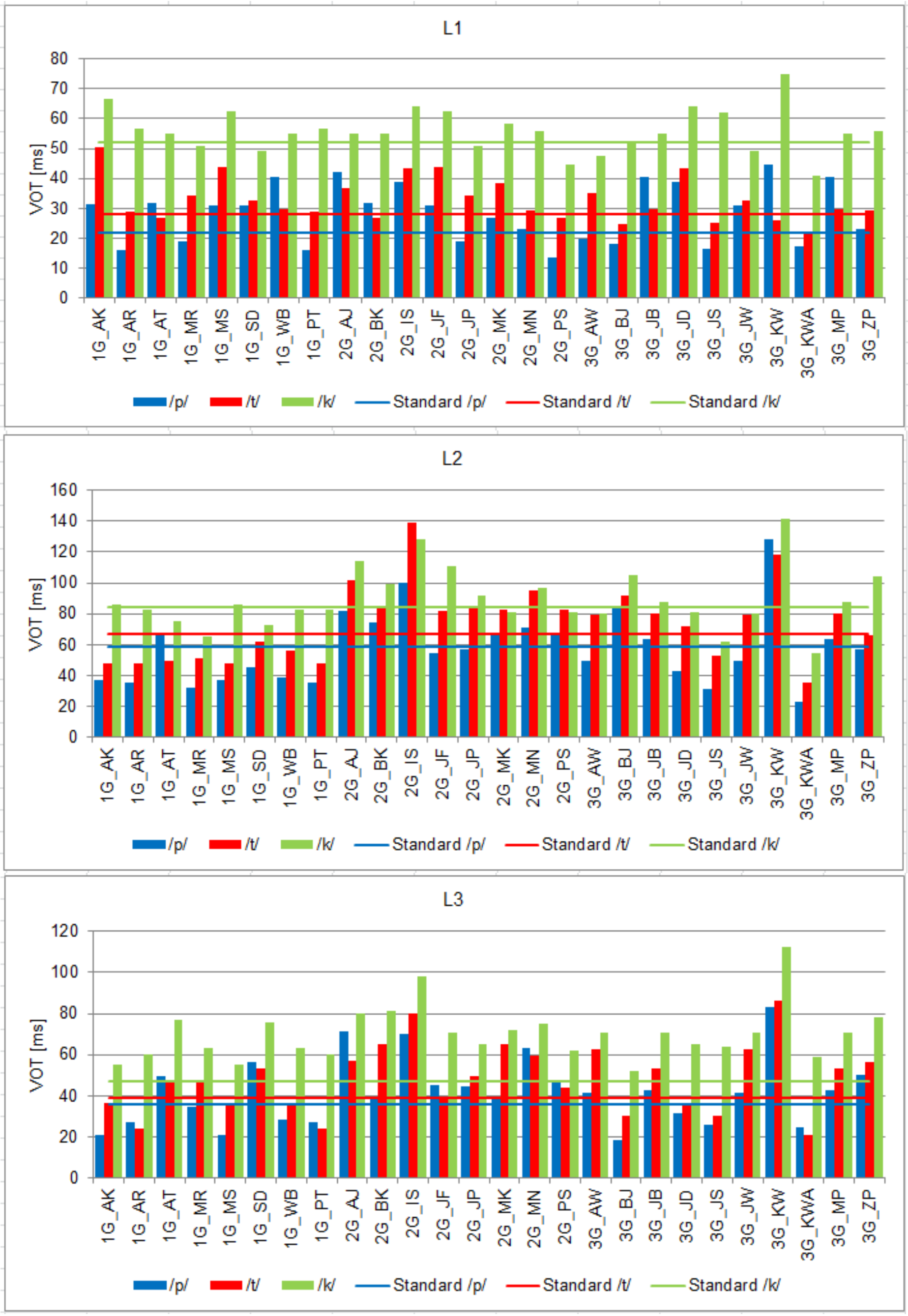
Language	N	Mean	SD	Median	Minimum	Maximum
L2 English	466	0.2	3.78	0.0	-8	15
L3 German	461	-2.2	2.86	-2.0	-8	7

5.4.2.8. Individual variation

The individual variation in the generated VOT values for /p/ /t/ /k/ was also explored as part of the general analysis in Group B. The following figures illustrate individual variation separately for L1 Polish, L2 English and L3 German against the selected reference VOT values for particular languages. As can be seen in Figures 48-50, the great majority of the participants followed the universal VOT pattern in all their language systems, with bilabials yielding the shortest VOT values, and velar plosives – the longest.

As expected, the VOT durations for L1 Polish exhibited less interspeaker variation than for the L2 or L3, and a fairly close correspondence to the VOT reference values for Polish. On the other hand, the individual VOT performance in L2 English appeared to be more variable, with some overshoots of the VOT English reference values (as in 2G_IS, 3G_KW) or undershoots (as in 1G_MR, 3G_KWA) as well as a number of individuals whose VOT values corresponded quite closely to the English norms (e.g. 2G_JP, 2G_JF, 2G_PS, 3G_MP, 3G_JB).

The greatest variability in the VOT distribution in Group B was visible in L3 German. As the selected reference values from the literature were more clumped together, very few, if any, individual distributions corresponded to this pattern. Interestingly, the observed most extreme departures from the German norm were represented by the same individuals as in the case of L2 English, i.e. 2G_IS, 3G_KW exemplifying VOT overshoots and 1G_MR, 3G_KWA representing undershoots. On the whole, the analysis of individual variation in L3 German seems to confirm the assumption that this language system is the least stable one, thus it demonstrates so much variability.



Figures 48-50. Individual variation in L1 Polish, L2 English and L3 German VOT for /p/ /t/ /k/ against the reference VOT values

5.4.2.9. Vocalic context effects

Non-parametric Kruskal-Wallis tests were used to investigate differences in the mean VOT with respect to the vocalic context in each of the participants' languages in Group B. The assumed alpha level was $\alpha=0.05$. The general assumptions were based on the universal tendencies, according to which the context of the high vowels (e.g. /i/) should generate longer VOT values in the preceding plosives than the context of the low vowels (e.g. /a/).

For the bilabial voiceless plosive /p/, the results of the performed Kruskal-Wallis Test pointed to statistically significant differences in the mean VOT values only in limited vocalic contexts and not in all the language systems. In L1 Polish, the only significant VOT durational difference was found for the mid vs. high vowel context, with the mid vowels /o, e/ preceding the initial plosives that generated significantly shorter VOT durations compared to the high vowel /i/ contexts. No statistically significant vocalic context effects related to the VOT durational differences were observed in L2 English. As far as L3 German is concerned, the high vowel context yielded significantly longer VOT than the low one (see Table 78).

Table 78. Comparison of mean VOT between /pa-, po/e-, pi-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	pa- vs po/e-	pa- vs pi-	po/e- vs pi-
L1	0.0111*	1.000000	0.051183	0.017065*
L2	0.2153	0.273232	0.637239	1.000000
L3	0.0034*	0.287220	0.002222*	0.262802

*p<0.05

The analysis of the vocalic context effects for the alveolar plosive /t/ demonstrated statistically significant differences in the mean VOT values in some contexts for all the language systems involved. Similar patterns were demonstrated by L1 Polish and L2 English, in which the low and mid vowel contexts generated shorter VOT durations compared to the high vowel /i/ contexts; however, there was no significant difference in the VOT measures between the low and mid vowels contexts for /t/. The vocalic effect differed, to some extent, in the case of L3 German, for which the high vowel context resulted in longer VOT compared to the mid vowel context, yet, the mid vowel context generated lower VOT measures than the low vowel context (see Table 79).

Table 79. Comparison of mean VOT between /ta-, to/e-, ti-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ta- vs to/e-	ta- vs ti-	to/e- vs ti-
L1	0.0000*	0.465354	0.000000*	0.000000*
L2	0.0109*	1.000000	0.029337*	0.025676*
L3	0.0001*	0.023164*	0.267133	0.000049*

*p<0.05

Finally, the results of the Kruskal-Wallis test for the multiple comparisons between the vocalic contexts for the velar plosive /k/ showed statistically significant differences in the mean VOT measures for all the languages involved; however, the relevant contexts differed. For L1 Polish, similarly to the previous results, the high /i/ vowel context resulted in longer VOT durations of the preceding consonant compared to both low and mid vowel contexts. A slightly different pattern was observed in L2 English, where significant differences in the VOT durations were found for the comparisons between the high and mid vowels as well as the low and mid vowels, in both cases the mid vowels generating shorter VOT durations. In L3 German, the only significant difference was found to hold between the high and mid vowel contexts, with the former resulting in longer VOT durations of the preceding velar plosive (see Table 80).

Table 80. Comparison of mean VOT between /ka-, ko/e-, ki-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ka- vs ko/e-	ka- vs ki-	ko/e- vs ki-
L1	0.0000*	0.089168	0.000000*	0.000000*
L2	0.0000*	0.000171*	1.000000	0.000344*
L3	0.0023*	0.445156	0.132559	0.001593*

*p<0.05

Recapitulating, the vocalic context effects observed in the data follow, to a large extent, the universal trends in the VOT durations, where the context of the high vowels results in longer VOT values in the preceding plosives than the context of the low vowels. The exhibited patterns varied between the language systems and different contexts, with L3 German demonstrating the lowest consistency in this respect, thus evidencing again the lowest stability of this language system.

Figures 51-53 present box plots of the mean VOT values for various vocalic contexts in the three respective languages in Group B.

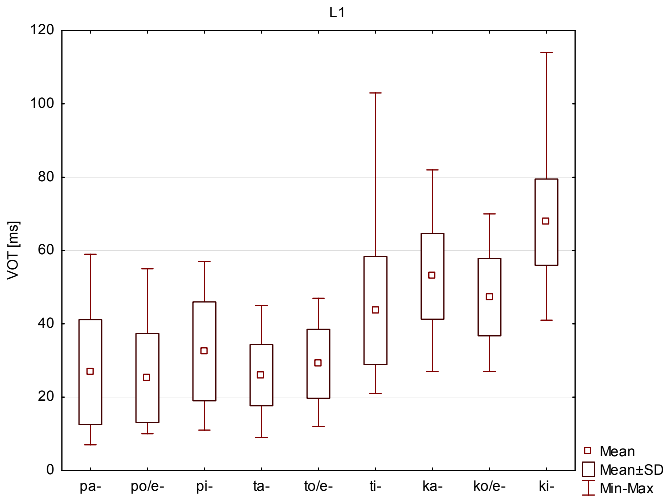


Figure 51. Box plot of mean VOT values for various vocalic contexts in L1 Polish

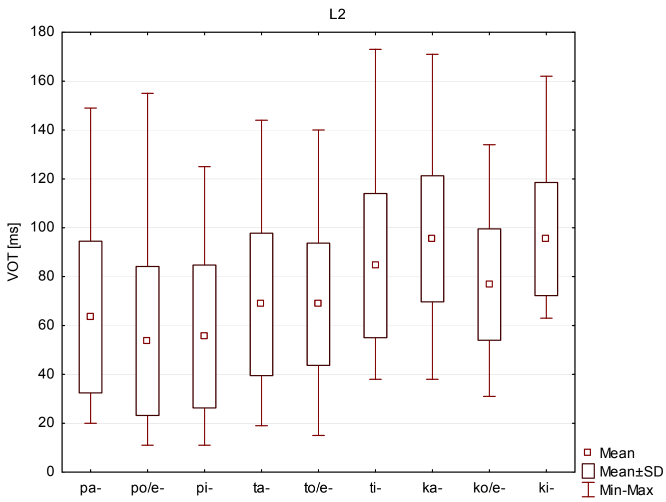


Figure 52. Box plot of mean VOT values for various vocalic contexts in L2 English

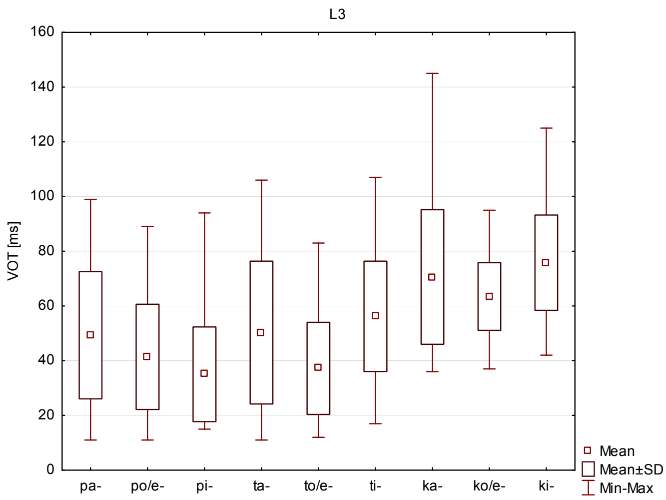


Figure 53. Box plot of mean VOT values for various vocalic contexts in L3 German

5.4.2.10. Analysis of variance

A two-factor ANOVA between the languages (L1, L2, L3) and the VOT durations of the voiceless plosive sounds /p, t, k/ was performed as part of the analysis of variance. Although the condition of normal distribution of the VOT variable in the studied subgroups was not met, the two factor analysis pointed to some existing interactions between the examined variables.

Table 81. Two-factor analysis of variance for /p, t, k/ and for L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	295.03 (0.05;2;1386)	0.000000*
plosive (/p/. /t/. /k/)	218.30 (0.05;2;1386)	0.000000*
Language*plosive	3.46 (0.05;4;1386)	0.008033*

*p<0.05

The findings demonstrate that the differences in the VOT values within the factors of the language (F (2; 1386) =295, p<.05) and the plosives (F (2; 1386) =218.3, p<.05) were found to be significant. Furthermore, the interaction between the languages and plosives on the VOT values was shown to be significant (F (4; 1386) =3.46, p<.05). It follows that there are significant differences between the VOT values for at least two languages, and for at least two plosives. The existing interaction between the languages and the plosives in Group B depends on the type of language (L1, L2, L3), as presented in Table 81 and Figure 54.

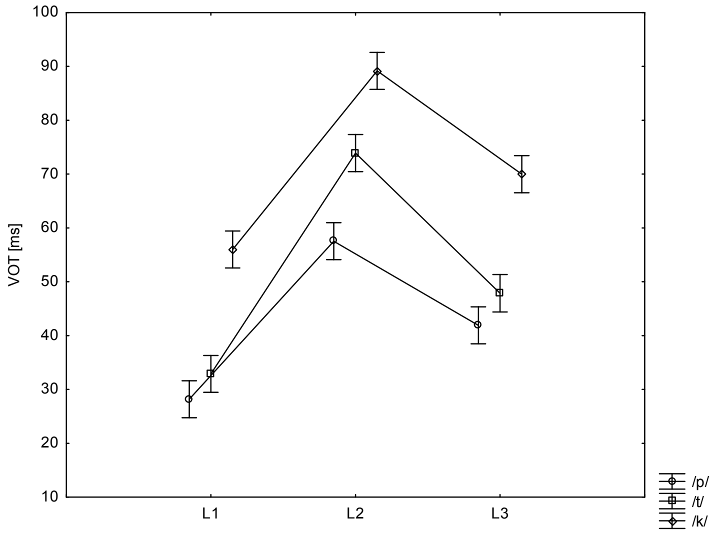


Figure 54. Interaction between the language and plosive factors

The following analysis intended to investigate the interaction of the vowel context and the language on the observed VOT durations in Group B. To this end, a two-factor ANOVA analysis was performed for the factors of the language (L1, L2, L3) and the vowel context following the voiceless plosives in the stressed onset positions in the target words (i.e. low _/a/, high _/i/, and mid _/e, o/). The results show that there are significant differences in the VOT values within the factor of the languages (F (2; 1386) =232.9, p<.05) and the vowel context (F (2; 1386) =25.48, p<.05). The findings indicate that significant differences between the VOT values were found for at least two languages, and for at least two vowel contexts. Furthermore, a significant interaction was found between the two factors (F (4; 1386) =3.43, p<.05) which depends on the type of the language (see Table 82 and Figure 55).

Table 82. Two-factor analysis of variance for vowel context /-a-/, /-o/e-/, /-i-/ and L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	232.94 (0.05;2;1386)	0.000000*
Vowel (/-a-/. /-o/e-/. /-i-/)	25.48 (0.05;2;1386)	0.000000*
Language*Vowel	3.43 (0.05;4;1386)	0.008392*

*p<0.05

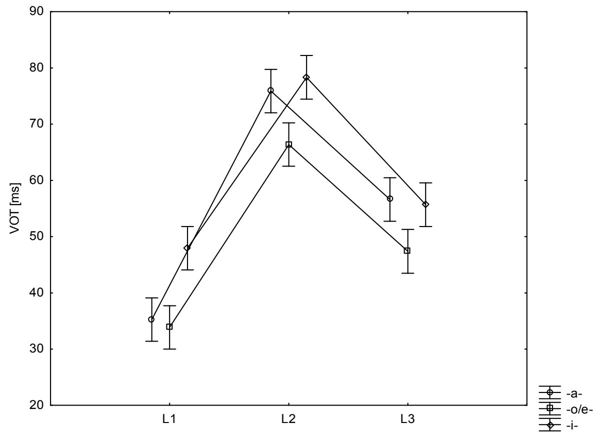


Figure 55. Interaction between the language and vowel context factors

The analysis of the patterns of interaction observed for the respective languages in Group B demonstrated that none of the language systems followed closely the universal patterns, according to which the longest VOT values are accompanied by the high vowel context /i/, medium for the mid vowels /e/, /o/, and the shortest for the low vowel context /a/. These regularities were partially observed in L1 Polish, L2 English and L3 German with the high vowel context generating on average the longest VOT values, whereas the mid and low vowel contexts yielded slightly different VOT duration patterns than the expected ones since the mid vowel contrasts /-o/e-/ were found to attract shorter VOT in the preceding plosive than the low vowel context /-a-/.

In order to investigate the potential effects of the proficiency level on the generated VOT patterns, the participants in Group B were subdivided into two proficiency subgroups, including the less advanced 1/2G group and the more advanced 3G group. The results of the two-factor ANOVA for the proficiency subgroups and for the L1, L2, L3 indicate that the differences in the VOT values within the factors of language (F (2; 1389)=219, p<.05) were significant; however, there were no significant differences between the mean VOT values with respect to the proficiency group (F (1; 1389) =1.35, p>.05). Further, the interaction between the factors of the language and the group on the mean VOT values in Group B was found to be significant (F (2; 1389) =5.32, p<.05) and to depend on the type of language (see Table 83 and Figure 56).

Table 83. Two-factor analysis of variance for proficiency groups and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	219.26 (0.05;2;1389)	0.000000*
Group (1/2G, 3G)	1.35 (0.05;1;1389)	0.245099
Language*Group	5.32 (0.05;2;1389)	0.004968*

*p<0.05

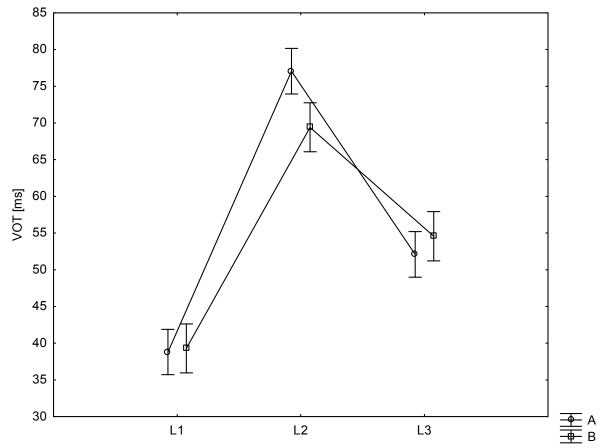


Figure 56. Interaction between the language and proficiency group factors (A:1/2G, B: 3G)

The conducted two-factor analysis of variance ANOVA for the proficiency groups and the place of articulation PoA /p, t, k/ separately for all the languages (L1, L2, L3) demonstrated the following results. For L1 Polish and L3 German, there were statistically significant differences between the mean VOT values for at least two places of articulation, i.e. for the L1(F (2; 462) =180.77, p<.05), and for the L3 (F (2; 455) =75.82, p<.05); there were no significant differences in VOT for the proficiency group; there was no significant interaction between the PoA and the proficiency groups.

For L2 English, there were statistically significant differences between the mean VOT values for at least two places of articulation (F (2; 460) =48.76, p<.05); there were statistically significant differences between the mean VOT values for the proficiency group (F (2; 460) =8.74, p<.05); there was no significant interaction between the PoA and the groups. The influence of the PoA on the mean VOT did not depend on the proficiency level of the group. The results are presented in Table 84 and Figure 57.

Table 84. Two-factor analysis of variance for proficiency groups and place of articulation PoA /p, t, k/ and for languages L1, L2, L3

Factor	L1		L2		L3	
	F	p	F	p	F	p
PoA (/p/, /t/, /k/)	180.77 (0.05;2;462)	0.000000*	48.76 (0.05;2;460)	0.000000*	75.82 (0.05;2;455)	0.000000*
Group (1/2G, 3G)	0.15 (0.05;1;462)	0.702645	8.74 (0.05;1;460)	0.003266*	1.63 (0.05;1;455)	0.202733
PoA*Group	0.34 (0.05;2;462)	0.712220	0.14 (0.05;2;460)	0.869907	0.03 (0.05;2;455)	0.973741

*p<0.05

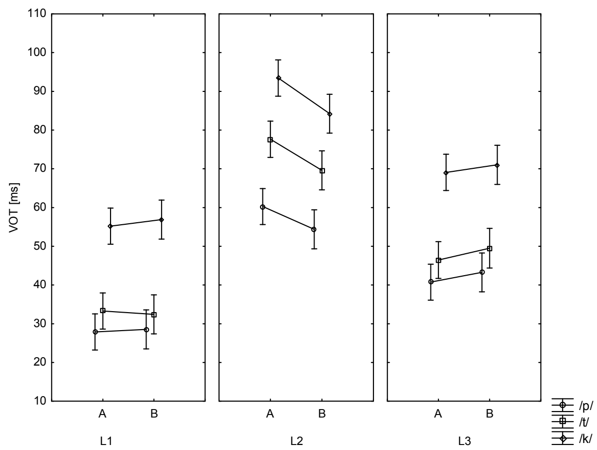


Figure 57. Interaction of PoA (/p/, /t/, /k/) and Group (A:1/2G, B: 3G) for languages L1, L2 and L3

5.4.2.11. Multiple regression analysis

The multiple regression analyses were performed to estimate the relationship between the dependent and independent variables in Group B. The first regression analysis investigated the relationship between the VOT values as the dependent variable and a number of independent variables including the language (L1, L2, L3), the place of articulation PoA (/p/, /t/, /k/), the vowel context and the proficiency group. The assumed levels of statistical significance pointed to a statistically significant interdependence between the VOT values and the following variables of the language, the place of articulation and the vowel contexts /-o/e-/ and /-i-/, whereas the influence of the proficiency group variable and the vowel /-a-/ was not found to be significant. The R squared result indicated that 45% of the variance model is accounted for by the independent variables.

A standardized coefficient, i.e. the beta coefficient was calculated to establish which of the independent variables had a greater effect on the dependent variable in the multiple regression analysis. The beta coefficient values suggest that Language 2, i.e. L2 (b=0.52), Language 1, i.e. L1 and L3 (b=-0.46) and PoA 1, i.e. /p/ (b=-0.36) had the biggest influence on the VOT values, whereas the impact of other variables such as the vowel context (Vowel 2 -i- , -o/e-) and PoA 2, i.e. /t/, /k/ is relatively smaller (see Tables 85 and 86).

Table 85. Multiple regression analysis 1

Variable	R	R^2	F	p
VOT	0.67	0.450	162.23 (0.05;7;1387)	0.000000*

*p<0.05

Table 86. Results of multiple regression 1 for variables

Effect	t	p	Beta
Target word	96.14	0.000000*	–
Language 1	-20.10	0.000000*	-0.461
Language 2	22.57	0.000000*	0.518
Group	1.36	0.174511	0.027
PoA 1	-15.75	0.000000*	-0.362
PoA 2	-4.56	0.000006*	-0.105

Vowel 1	0.82	0.412022	0.019
Vowel 2	-7.48	0.000000*	-0.172

*p<0.05

A further multiple regression analysis was performed for the L3 VOT values as the dependent variable and the L1 VOT and L2 VOT as the independent variables in Group B. The analysis showed a statistically significant interdependence among the variables under investigation. The regression model is explained in 42.5% by the independent variables L1 VOT and L2 VOT (R^2=0.425). The bigger impact on the dependent variable L3 German VOT is exerted by the L2 English VOT values (b=0.5) than the L1 Polish VOT (b=0.25) (see Tables 87 and 88).

Table 87. Multiple regression analysis 2

Variable	R	R^2	F	p
L3 VOT	0.65	0.425	168.34 (0.05;2;456)	0.000000*

*p<0.05

Table 88. Results of multiple regression 2 for variables

Effect	t	p	Beta
Target word	4.72	0.000003*	
L1 VOT	6.43	0.000000*	0.253
L2 VOT	12.77	0.000000*	0.502

*p<0.05

A final regression analysis was calculated for the L3 VOT values as the dependent variable and other independent variables including the L1 VOT, L2 VOT, the place of articulation /p, t, k/, the vowel context and the proficiency group. The regression model is explained in 48.2% by the independent variables. The assumed level of testing probability points to a statistically significant impact of such independent variables as the L1 VOT, the L2 VOT, the PoA, the vowel contexts /-a-/, /-i-/ and the group, with the only insignificant variable being the vowel context /-o/e-/. The greatest impact on the dependent variable L3 VOT is exerted by the L2 VOT (b=0.48), whereas it is considerably smaller for other independent variables such as PoA 1 and PoA 2, the proficiency group, L1 VOT and the vowel context 1/-a-/, /-i-/ (see Tables 89 and 90).

Table 89. Multiple regression analysis 3

Factor	R	R^2	F	p
L3 VOT	0.69	0.482	59.95 (0.05;7;451)	0.000000*

*p<0.05

Table 90. Results of multiple regression 3 for variables

Effect	t	p	Beta
Target word	6.39	0.000000*	
VOT L1	2.01	0.044820*	0.105
VOT L2	12.26	0.000000*	0.482
Group	-3.31	0.001008*	-0.113
PoA 1	-2.77	0.005755*	-0.132
PoA 2	-3.68	0.000258*	-0.153
Vowel 1	2.52	0.012243*	0.101
Vowel 2	-1.91	0.056827	-0.078

*p<0.05

5.4.3. Results for Group C

In the following sections, the results for Group C with L1 Polish, L2 German and L3 English will be discussed.

5.4.3.1. Mean VOT values for L1, L2 and L3

The Shapiro–Wilk test of normality was administered to verify the null hypothesis for this test that the VOT measurements data are normally distributed. Since the chosen alpha level was 0.05 and the p-values for the VOT variable in several categories were less than 0.05, then the null hypothesis that the data are normally distributed had to be rejected. Consequently, non-parametric tests were applied for further statistical analyses of the data in Group C.

Tables 91-93 present the mean results of the VOT measurements of the voiceless plosives /p/, /t/, /k/ for the target words read in the carrier phrases in the participants' L1 Polish, L2 German and L3 English.

Table 91. Mean VOT values for target words in L1 Polish (ms)

Words in L1	N	Mean	SD	Median
pan	33	21.3	11.0	19.0
para	33	19.4	6.0	18.0
pech	33	19.0	12.6	16.0
pole	33	28.1	12.0	25.0
pik	33	22.9	9.4	21.0
piwo	33	28.9	10.6	27.0
tak	33	23.1	9.0	20.0
tango	33	20.2	6.2	20.0
tik	62	33.5	11.8	32.0
testy	33	24.0	8.6	20.0
tiry	33	53.6	9.7	54.0
kat	33	47.7	14.8	47.0
kanon	33	49.6	13.0	51.0
kot	33	45.5	10.6	45.0
kelner	33	48.1	12.2	49.0
kit	32	60.4	11.8	62.0
kino	31	62.9	12.6	65.0

Table 92. Mean VOT values for target words in L2 German

Words in L2	N	Mean	SD	Median
Paar	32	50.3	16.9	50.0
Panne	33	41.4	19.6	36.0
Pelz	33	31.6	15.6	28.0
Pose	33	43.8	17.0	41.0
Pirsch	30	33.0	10.7	31.0
Pinsel	33	29.7	12.8	28.0
Tanz	33	34.9	18.9	27.0
Tasse	33	34.9	18.8	28.0
Text	33	27.7	13.3	25.0
Tochter	33	30.8	11.4	29.0
tief	32	59.8	16.0	58.0
ticken	31	40.5	14.7	36.0
Kalk	33	63.0	13.2	61.0
Karte	33	59.7	11.7	60.0
kommen	33	68.3	16.4	69.0
Kegel	33	65.5	20.3	65.0
Kind	33	67.1	17.0	64.0
Kilo	31	86.4	13.5	85.0

Table 93. Mean VOT values for target words in L3 English

Words in L3	N	Mean	SD	Median
pan	32	38.4	19.0	34.5
party	33	36.9	17.5	37.0
pot	32	39.0	16.4	33.0
pencil	33	24.9	9.9	24.0
peace	33	32.8	16.2	27.0
Peter	32	31.4	16.2	27.5
task	33	36.7	16.9	34.0
taxi	33	30.4	14.3	27.0
test	33	34.4	15.8	28.0
toffee	33	35.5	16.6	30.0
team	33	62.0	21.2	65.0
teacher	31	46.6	11.9	45.0
cat	31	64.0	19.6	60.0
carpark	32	65.3	17.0	66.5
kept	33	53.7	17.0	49.0
coffee	33	51.7	15.9	48.0
keen	31	87.1	16.7	89.0
keeper	32	71.7	17.8	72.0

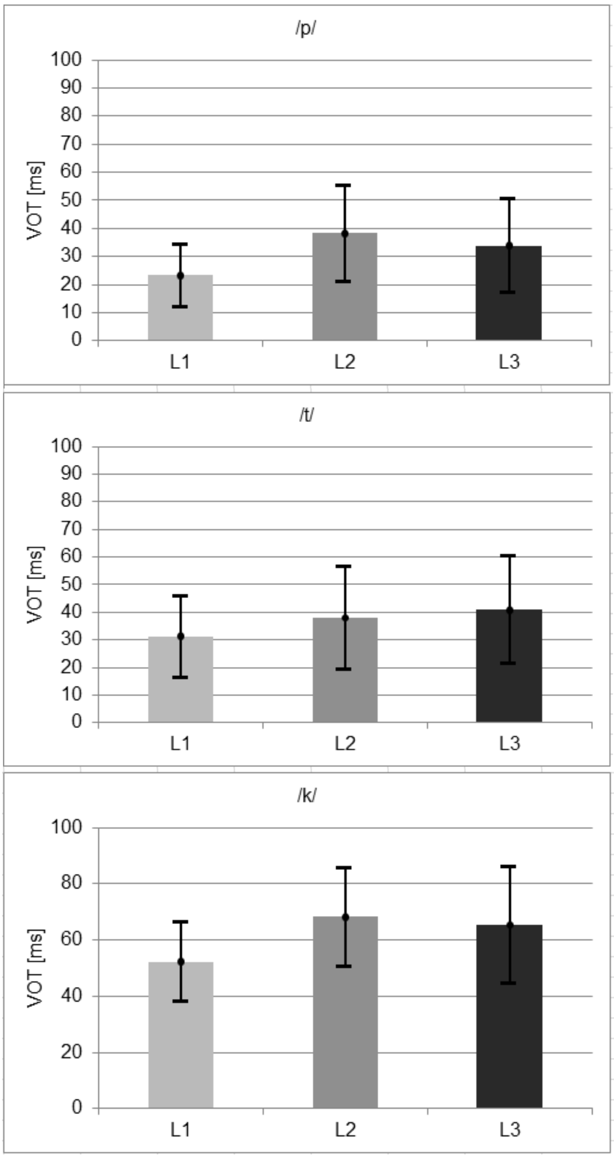
The results of the acoustic measurements of the mean voice onset time for the voiceless plosives /p/, /t/, /k/ in the stress onset positions in the participants' L1 Polish, L2 German and L3 English are presented in Table 94 and Figures 58-60.

The conducted durational measurements in Group C demonstrate that the multilingual participants produced the voiceless plosives in the stress onset positions with the mean voice onset time values that were the shortest in L1 Polish (/p/=23 ms, /t/=31 ms, /k/=52 ms), and had a somewhat longer-lag both in their L2 German (/p/=38 ms, /t/=38 ms, /k/=68 ms) and L3 English (/p/=34 ms, /t/=41 ms, /k/=65 ms).

Table 94. Mean VOT values for L1, L2 and L3

VOT	Language	N	Mean	SD	Median
/p/	L1	198	23.3	11.1	21.0
	L2	194	38.3	17.2	35.5
	L3	195	33.9	16.6	29.0
/t/	L1	194	31.3	14.8	26.0
	L2	195	38.0	18.7	33.0
	L3	196	40.9	19.4	37.0

/k/	L1	195	52.2	14.1	51.0
	L2	196	68.2	17.6	68.0
	L3	192	65.3	20.8	63.5



Figures 58-60. VOT measurements for /p, t, k/ in L1 Polish, L2 German, L3 English

5.4.3.2. Cross-language comparison of VOT means

The first series of the statistical tests was conducted in order to investigate the language effect in the VOT measurements in Group C. A non-parametric Kruskal-Wallis ANOVA was used to compare the mean VOT values for /p/ /t/ /k/ across the languages of the multilingual participants. The assumed alpha level was 0.05.

The performed ANOVA pointed to significantly different values for all the initial voiceless plosives between L1 Polish and L2 German as well as between L1 Polish and L3 English ($p<.05$), however, the differences between the mean VOT values for two out of three plosives in L2 German and L3 English were not found to be statistically significant. A detailed analysis of the cross-language comparison of the VOT means is presented in Table 95.

Table 95. Mean VOT comparison for /p t k/ in L1 Polish, L2 German, L3 English

Kruskal-Wallis Test	p for multiple comparisons		
	L1 vs L2	L1 vs L3	L2 vs L3
/p/	0.000000*	0.000000*	0.018004*
/t/	0.000901*	0.000001*	0.365420
/k/	0.000000*	0.000000*	0.193789

*p<0.05

As far as the measurements for /p/ are concerned, the *p* values for the multiple comparisons point to statistically significant differences in the mean VOT for this bilabial voiceless plosive between all the languages involved, i.e. between the L1 and the L2 (the mean VOT values in L1 Polish are lower than in L2 German); between the L1 and the L3 (the mean VOT values in L1 Polish are lower than in L3 English); between the L2 and the L3 (the mean VOT values in L2 German are higher than in L3 English).

In the case of the mean VOT measurements for /t/, the *p* values for the multiple comparisons also demonstrated statistically significant differences between the languages involved, i.e. between the L1 and the L2 (the mean VOT values in L1 Polish are lower than in L2 German); between the L1 and L3 (the mean VOT values in L1 Polish are lower than in L3 English). However, the difference between the mean VOT values for /t/ in L2 German and L3 English did not prove to be significant.

Similar cross-linguistic patterns were observed also for the /k/ measurements. The performed multiple comparisons also demonstrated statistically significant differences between the languages involved for /k/, i.e. between the L1 and the L2 (the mean VOT values in L1 Polish are lower than in L2 German); between the L1 and the L3 (the mean VOT values in L1 Polish are lower than in L3 English). However, the difference between the mean VOT values for /k/ in L2 German and L3 English did not prove to be significant as in the case of /t/.

Recapitulating, the performed tests yielded the cross-linguistic differences that proved to be significant between the native tongue (i.e. Polish) and both non-native languages, however, the pairwise comparison of means between L2 German and L3 English demonstrated a significant difference only for the bilabial /p/ ($p<0.05$), whereas the differences for the alveolar and velar plosives /t/ and /k/ were found to be non-significant in Group C.

The following box plots (Figures 61-63) illustrate the observed tendencies in the VOT patterns in Group C, in the respective languages separately for the stressed onset plosives /p/ /t/ and /k/. It is interesting to observe that the distribution in L2 German and L3 English shows similar trends with respect to the VOT means, the standard deviation as well as the minimum-maximum range, whereas the mean values for L1 Polish remain significantly lower and the minimum-maximum range is less pronounced.

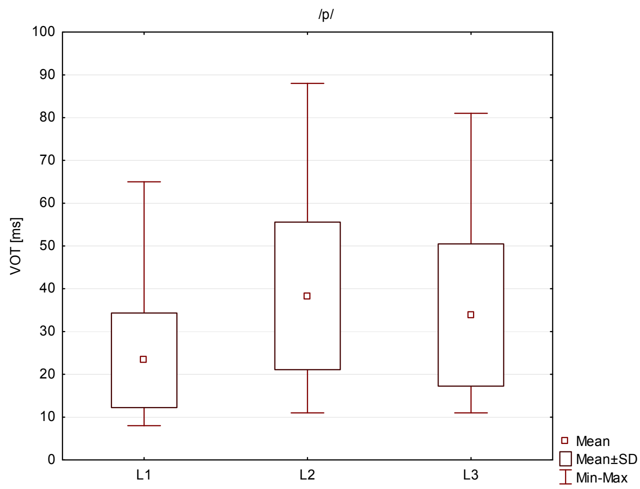


Figure 61. Box plot of mean VOT values for /p/ in L1 Polish, L2 German and L3 English

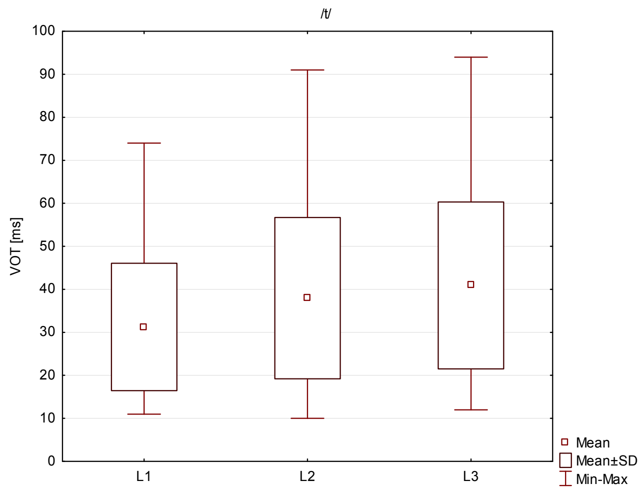


Figure 62. Box plot of mean VOT values for /t/ in L1 Polish, L2 German and L3 English

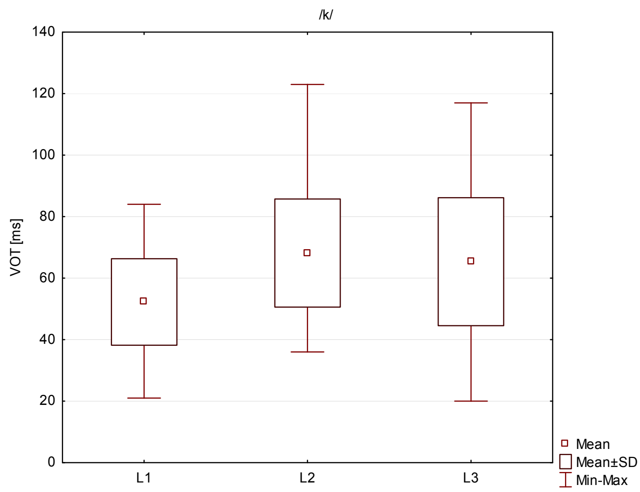


Figure 63. Box plot of mean VOT values for /k/ in L1 Polish, L2 German and L3 English

The language effect was thus observed to hold between the first language and the two non-native phonological systems. The mean voice onset time values in L1 Polish were considerably lower than the respective values in L2 German and L3 English, which, on the other hand, displayed quite similar patterns of distribution in Group C.

5.4.3.3. Proficiency group effect

In order to examine the language proficiency group effect on the VOT acquisition in Group C, the non-parametric Mann-Whitney U tests were performed for the mean VOT values for /p/, /t/, /k/ in L1, L2 and L3 for the subgroup 1EG (less proficient in L3 English, N=15) vs. the 2/3/4 EG subgroup (more advanced in L3 English, N=18) (see Table 96).

Table 96. Results of Mann-Whitney test between two groups (lower proficiency 1EG vs. higher proficiency 2/3/4EG) for /p, t, k/ in L1, L2 and L3

Sounds	Language	Z corr.	p
/p/	L1	1.76	0.078109
	L2	-1.12	0.264079
	L3	3.51	0.000444*
/t/	L1	-2.06	0.039703*
	L2	-2.00	0.045640*
	L3	1.30	0.193478
/k/	L1	-0.43	0.668210
	L2	-2.69	0.007048*
	L3	1.39	0.163811

*p<0.05

The results of the analyses showed significant differences between the VOT values with respect to the language proficiency group in the following conditions:

– /p/ in L3 English – the mean VOT in the lower proficiency group 1EG is higher than in the 2/3/4EG group
– /t/ in L1 Polish – the mean VOT in the 1EG group is lower than in the 2/3/4EG group
– /t/ in L2 German – the mean VOT in the 1EG group is lower than in the 2/3/4EG group
– /k/ in L2 German – the mean VOT in the 1EG group is lower than in the 2/3/4EG group.

Summing up, as far as the proficiency group effect is concerned, no major consistent regularities were observed, apart from the tendency in L2 German for longer VOT values in the more advanced group.

The following box plots (Figures 64-66) illustrate the observed tendencies in the group comparison based on L3 the proficiency criteria (1EG – the less advanced, 2/3/4EG – the more advanced group).

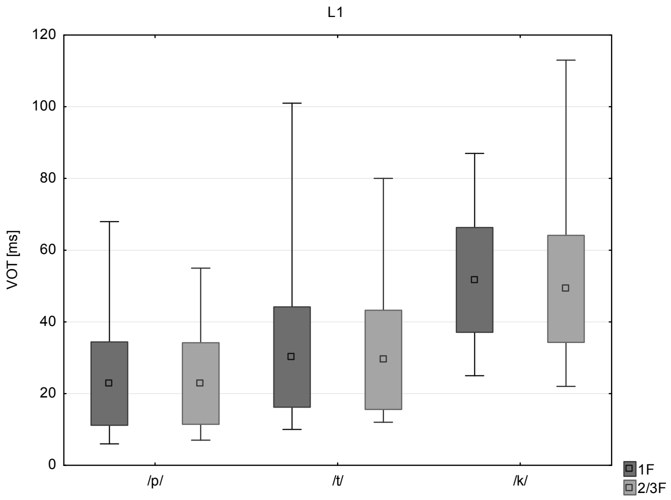


Figure 64. Box plot of mean VOT values for /p/, /t/, /k/ in L1 Polish – group comparison

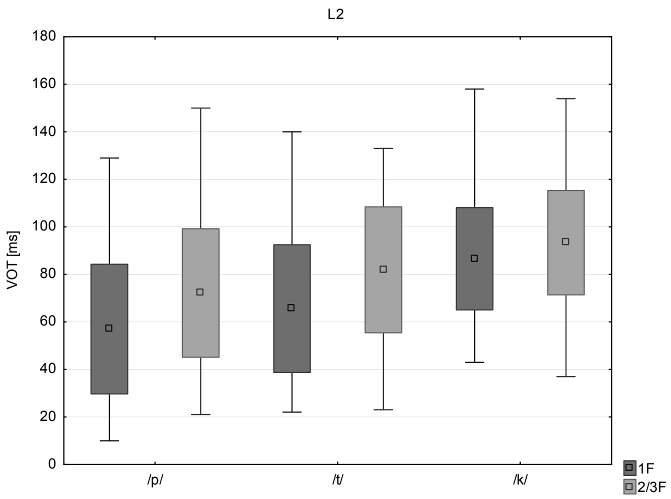


Figure 65. Box plot of mean VOT values for /p/, /t/, /k/ in L2 German – group comparison

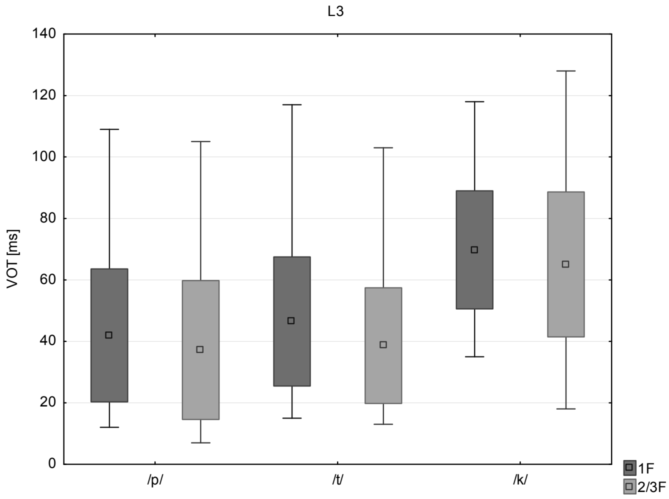


Figure 66. Box plot of mean VOT values for /p/, /t/, /k/ in L3 English – group comparison

5.4.3.4. Cross-linguistic correlations between VOT values

In order to investigate the dependence between two selected variables in Group C, i.e. the observed VOT values in L3 English and those of the native L1 Polish, as well as between L3 English and L2 German, the Pearson product-moment correlation analysis was applied (α=0.05).The calculated coefficients pointed to prevailingly positive weak correlations between the mean VOT values.

In the case of the voiceless bilabial plosive /p/, there was only a statistically significant correlation for one pair of variables, i.e. L2 German and L3 English. It indicated a weak positive correlation (R=0.28) between the non-native languages in the VOT values for /p/. There was no statistically significant correlation between L1 Polish and L3 English in this respect (see Table 97).

Table 97. Pearson's correlation for /p/

Pair of variables for /p/	N	R	t	p
VOT L1 & VOT L3	191	-0.04	-0.59	0.553991
VOT L2 & VOT L3	191	0.28	3.99	0.000095*

*p<0.05

As far as the Pearson's correlations for /t/ are concerned, a weak positive correlation was found between L1 Polish and L3 English (R=0.25). In the case of the second pair of variables, i.e. L2 German and L3 English, the coefficient was significantly higher and pointed to a moderate correlation between the non-native languages (R=0.5) (see Table 98). This correlation is illustrated in Figure 67 on the presented scatterplot.

Table 98. Pearson's correlation for /t/

Pair of variables for /t/	N	R	t	p
VOT L1 & VOT L3	189	0.25	3.51	0.000566*
VOT L2 & VOT L3	189	0.50	7.86	0.000000*

*p<0.05

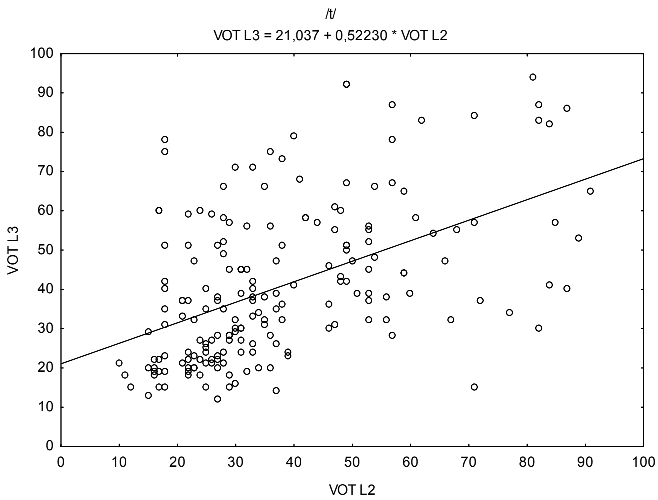


Figure 67. Scatterplot of VOT for /t/ between L2 German and L3 English

In the case of the velar plosive /k/, the obtained Pearson's correlations for both pairs of variables were in the weak range; with the coefficient ranging from R=0.31 for L1 Polish and L3 English to R=0.25 for L2 German and L3 English (see Table 99).

Table 99. Pearson's correlation for /k/

Pair of variables for /k/	N	R	t	p
VOT L1 & VOT L3	188	0.31	4.51	0.000012*
VOT L2 & VOT L3	188	0.25	3.57	0.000446*

*p<0.05

It appears impossible to state unequivocally whether the L3 VOT values were correlated more with the native values or those of another foreign language, as the differences between the coefficients were relatively small. There is a tendency for some more numerous and slightly stronger correlations between the L2 German and L3 English VOT values when compared to the ones between L1 Polish and L3 English, however, the covariance between the respective variables remains rather weak in nature in Group C.

5.4.3.5. Comparison to VOT reference values

In order to compare the acoustic measurements obtained in Group C to the VOT reference values from the literature for the respective languages, a series of analyses was performed. The one-sample t-tests and a non-parametric Wilcoxon signed-rank matched pairs tests were administered to compare the calculated mean VOT durations for /p, t, k/ in L1 Polish, L2 German and L3 English to the monolingual reference values often quoted in the literature, i.e. Keating et al. 1981 for Polish, Angelowa and Pompino-Marschall 1985 for German and Lisker and Abramson 1964 for English.

Table 100. Comparison to VOT reference values in Polish, German and English ([1]Keating et al., 1981; [2]Angelowa and Pompino-Marschall, 1985; [3]Lisker and Abramson, 1964)

Language	Parameter	VOT		
		/p/	/t/	/k/
Polish	Ref. VOT[1]	22	28	52
	L1 Mean	23.3	31.3	52.2
	SD	11.1	14.8	14.1
	p	0.105834	**0.002542***	0.815050
German	Ref. VOT[2]	36	39	47
	L2 Mean	38.3	38.0	68.2
	SD	17.2	18.7	17.6
	p	0.061432	0.436708	**0.000000***
English	Ref. VOT[3]	59	67	84
	L3 Mean	33.9	40.9	65.3
	SD	16.6	19.4	20.8
	p	**0.000000***	**0.000000***	**0.000000***

*p<0.05

The results demonstrated that the VOT measurements for L1 Polish and L2 German did not differ significantly from the monolingual native norms as reported in the literature, whereas the VOT values for L3 English did not correspond to the reported norms for English (see Table 100). More specifically, the VOT values for the voiceless stops in L1 Polish were very close to the reported monolingual norms (cf. Keating et al., 1981) in the case of /p/ and /k/ and slightly higher for /t/ (i.e. 31.3 vs. 28 ms), although still within the accepted 5-10 ms range.

Similarly in L2 German, the reported values for /p/ and /t/ corresponded closely to the reference values with the exception of the velar plosive /k/ which exhibited a longer mean duration of voice onset time than the monolingual reference, i.e. 68.2 vs. 47 ms (cf. Angelowa and Pompino-Marschall, 1985). However, the VOT measurements for L3 English were found to be significantly different from the monolingual norms (cf. Lisker and Abramson, 1964). The bilabial, alveolar and velar voiceless plosives were realized, on average, with a shorter lag than the reference values (/p/ 33.9 vs. 59 ms; /t/ 40.9 vs. 67 ms; /k/ 65.3 vs. 84 ms).

All in all, the observed VOT values of the multilingual participants in Group C did not depart significantly from the reference norms in the case of L1 Polish and L2 German (except for the velar /k/ which was found to have a longer lag). Conversely, a considerable VOT shortening was also observed for L3 English when compared to the literature reference values. The English stops were implemented by the multilingual participants with a shorter lag and thus the L3 phonetic norms were not approximated successfully. An interesting observation is that the L3 VOT values approximated, to a large extent, the L2 values, which may be interpreted as the influence from the second language system rather than the mother tongue.

5.4.3.6. Comparison to control groups

To verify further the reference values presented in the literature, the VOT measurements in Group C were compared to those generated from the control groups involved in the study. The German control group consisted of 17 native Germans (i.e. students and lecturers at the University of Leipzig), who were recorded reading the word list in the carrier phrases for German. The English control group consisted of 17 native English speakers (i.e. lecturers at the University of Essex and Adam Mickiewicz University in Poznań), who were recorded reading the list for English.

The generated control mean values were /p/=62, /t/=67, /k/=82 ms for German and /p/=61, /t/=75, /k/=79 ms for English. In the case of the English control group, the values were in the range of reference values from the literature; however, the German control group measurements exceeded considerably the reference values for German from the literature (see Table 111).

The one-sample t-tests and a non-parametric Wilcoxon signed-rank matched pairs tests were performed to compare the observed mean VOT durations for /p, t, k/ in the participants' L2 German and L3 English to the respective control groups (α=0.05).

Table 111. Comparison of mean VOT for /p, t, k/ in L2 German and L3 English to Control Groups values

Language	Parameter	VOT		
		/p/	/t/	/k/
German	German Controls	62	67	82
	L2 Mean	38,3	38,0	68,2
	SD	17,2	18,7	17,6
	p	0,000000*	0,000000*	0,000000*
English	English Controls	61	75	79
	L3 Mean	33,9	40,9	65,3
	SD	16,6	19,4	20,8
	p	0,000000*	0,000000*	0,000000*

*p<0.05

Statistically significant differences were found for the comparisons of the VOT durations of the voiceless plosives in Group C to the control native groups, with the latter exhibiting longer lag values than those in the participants' L2 German or L3 English. A potential explanation may be that when compared to the control group target values, the multilingual participants do not approximate fully these targets either in their L2 German or L3 English, producing instead some compromise values that may be intermediate between their L1 and the L2/L3 target norms.

5.4.3.7. VOT goodness of fit

In order to evaluate the participants' approximation to the target control values in Group C, the same procedure was followed as in Groups A and B. The calculation of the measure of the nativeness effect or VOT goodness of fit, based on the degree of approximation or deviation from the control VOT values, is explained in section 2.1.7. Consequently, the participants' measurements were assigned to particular categories, the higher the number the more they differed from the control baseline durations, with '+ values' indicating a longer lag and '– values' shorter VOT durations. Table 112 shows the distribution of percentage scores across the approximation threshold levels for L2 German and L3 English in comparison to the respective control VOT durations.

Table 112. Level of approximation of L2 and L3 VOT values to control VOT (in % of participants)

Approximation level	L2 German to Control VOT			L3 English to Control VOT		
	/p/	/t/	/k/	/p/	/t/	/k/
-8	1.5	2.1	0	2.6	1.5	0
-7	8.2	11.3	0	14.9	18.4	1.0
-6	12.4	16.4	0	17.4	13.8	1.0
-5	20.6	24.1	4.6	17.9	17.3	7.8
-4	10.8	12.8	10.2	11.8	11.7	12.5
-3	14.9	3.6	18.4	9.2	11.2	17.7
-2	8.2	11.3	12.2	9.2	9.2	9.9
-1	5.2	6.7	15.8	4.6	6.6	10.9
0	10.3	5.1	26.0	5.6	5.6	23.4
1	3.6	0.5	7.7	3.6	3.1	6.8
2	2.6	5.1	3.1	1.0	1.5	3.1
3	0.5	1.0	1.5	2.1	0	3.1
4	1.0	0	0.0	0	0	2.6
5	0	0	0.5	0	0	0

Table 113. Approximation level mean results

Language	N	Mean	SD	Median	Minimum	Maximum
L2 German	585	-2.9	2.6	-3.0	-8	5
L3 English	583	-3.2	2.7	-4.0	-8	4

Summing up, Table 13 presents the mean results of the approximation level analysis pointing to the comparable scores for L2 German (M=-2.9) and L3 English (M=-3.2). It appears that, on average, the L1 nativeness effect based on the VOT approximation to the respective control values was comparable for both non-native languages, with L3 English VOT durations being slightly more deviant from the baseline values.

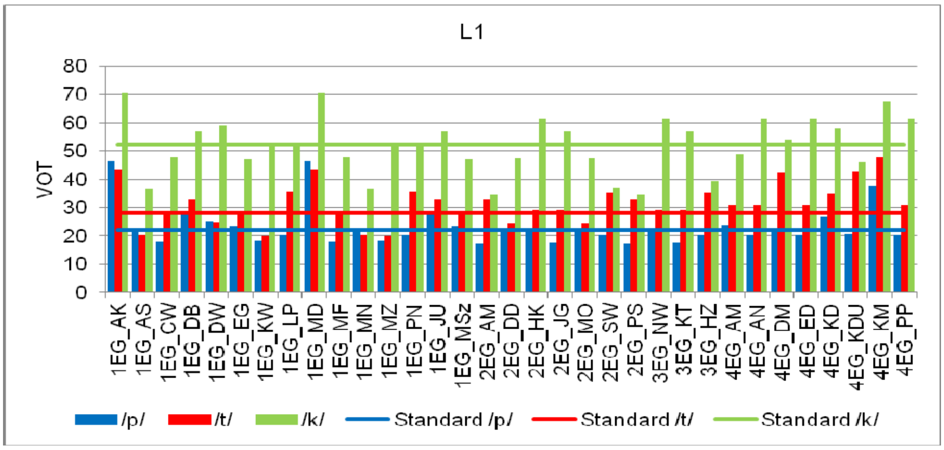
5.4.3.8. Individual variation

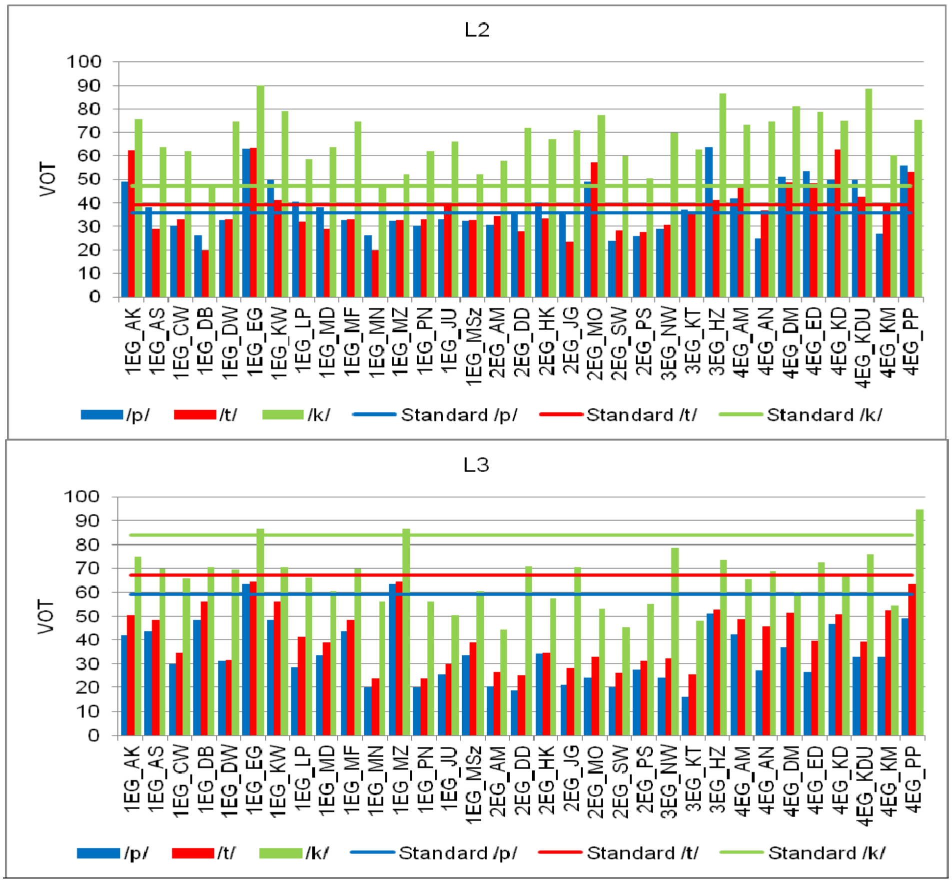
The analysis of the VOT measurements investigated also the individual variation in the generated VOT values for /p/ /t/ /k/ in Group C. The following figures illustrate individual variation separately for the L1, L2 and L3 against the selected reference VOT values for particular languages.

The greatest variability in the VOT distribution is visible in the case of L3 English. As can be seen in Figures 68-70, all of the participants followed the universal VOT patterns, with the bilabials yielding the shortest VOT values, and the velar plosives the longest.

On the whole, the observed L3 values are undershoots of the reference VOT measurements, with such individuals as 2EG_AM, 2AM_SW, 3EG_KT representing the most extreme departures from the norm. On the other hand, the L3 performance of some individual participants like 1EG_EG, 1EG_MZ, 4EG_PP appears to be fairly close to the target English VOT values.

As expected, the VOT durations for L1 Polish exhibited less interspeaker variation and fairly close correspondence to the VOT reference values for Polish. On the other hand, the individual variability in L2 German appeared intermediate between the other two language systems, i.e. it was more variable than the native language but less extreme than in the case of the other, later acquired, foreign language, i.e. the L3.





Figures 68-70. Individual variation in L1 Polish, L2 German L3 English VOT for /p/ /t/ /k/ against the reference VOT values

5.4.3.9. Vocalic context effects

In order to investigate the differences in the mean VOT with respect to the vocalic context in each of the participants' languages in Group C, a non-parametric Kruskal-Wallis test was applied (the assumed alpha level was α=0.05).

The results of the performed test for the bilabial voiceless plosive /p/ point to statistically significant differences in the mean VOT values between /pa-, po/e-, pi-/ contrasts only for L1 Polish and L2 German and only in limited contexts. In L1 Polish, the low and mid vowels preceding the initial plosives generated shorter VOT durations compared to the high vowel contexts. In L2 German, the only context that demonstrated significant differences was the one between the high and low vowels. The observed patterns are in line with the universal tendencies in the vocalic context effects on the VOT durations. No statistically significant differences in the VOT durations resulting from the vocalic context were found in L3 English (see Table 114).

Table 114. Comparison of mean VOT between /pa-, po/e-, pi-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	pa- vs po/e-	pa- vs pi-	po/e- vs pi-
L1	0.001500*	0.892232	0.001354*	0.040987*
L2	0.000100*	0.056804	0.000051*	0.139978
L3	0.110400	0.227043	0.190544	1.000000

*p<0.05

The analysis of the vocalic context effects for the alveolar plosive /t/ demonstrated statistically significant differences in the mean VOT values for L1 Polish, L2 German and L3 English. As expected, the high vowel /i/ contexts resulted in longer VOT durations compared to the low vowel /a/ and the mid vowels /o, e/ in all the languages. Additionally, L1 Polish, unlike the L2 and the L3, demonstrated statistically significant differences between the mid and low vowel contexts, with the latter contexts generating shorter VOT durations (see Table 115).

Table 115. Comparison of mean VOT between /ta-, to/e-, ti-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ta- vs to/e-	ta- vs ti-	to/e- vs ti-
L1	0.000000*	0.004793*	0.000000*	0.000001*
L2	0.000000*	0.438186	0.000000*	0.000000*
L3	0.000000*	1.000000	0.000000*	0.000000*

*p<0.05

The results of the Kruskal-Wallis test for the multiple comparisons between the vocalic contexts for the velar plosive /k/ showed statistically significant differences in the mean VOT measures for all the languages involved. Consistently with the previous results for /t/ in L1 Polish, L2 German and L3 English, the high /i/ vowel context resulted in longer VOT durations of the preceding consonant /k/, compared to both low and mid vowel contexts. Additionally, in L3 English significant differences in

VOT durations were found for the comparisons between the low and mid vowel contexts, with the former generating longer VOT durations in the preceding plosive (see Table 116).

Table 116. Comparison of mean VOT between /ka-, ko/e-, ki-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ka- vs ko/e-	ka- vs ki-	ko/e- vs ki-
L1	0.000000*	1.000000	0.000000*	0.000000*
L2	0.000000*	0.243550	0.000004*	0.006098*
L3	0.000000*	0.002272*	0.000487*	0.000000*

*p<0.05

All in all, the vocalic context effects observed in the data follow, to a considerable extent, the universal trends with the high vowel contexts generating longer VOT durations of the preceding plosive consonants than the low or mid vowel contexts. However, L1 Polish and L2 German exhibited somewhat more consistency in these patterns, whereas L3 English was a bit less consistent in this respect, possibly indicating a lower stability of the phonological system.

See Figures 71-73 presenting the box plots of the mean VOT values for various vocalic contexts in the three respective languages in Group C.

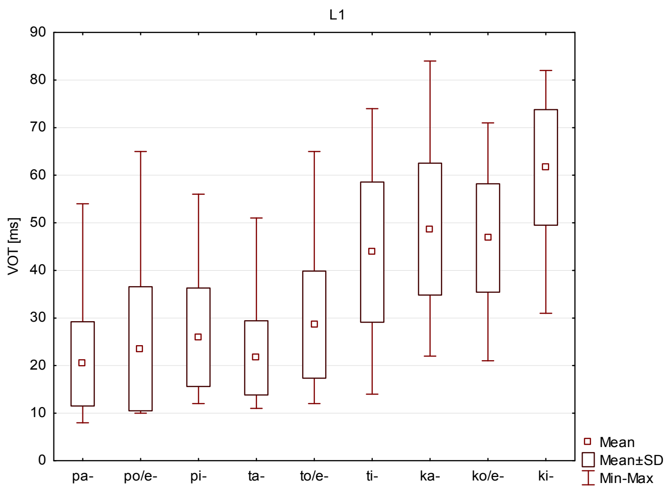


Figure 71. Box plot of mean VOT values for various vocalic contexts in L1 Polish

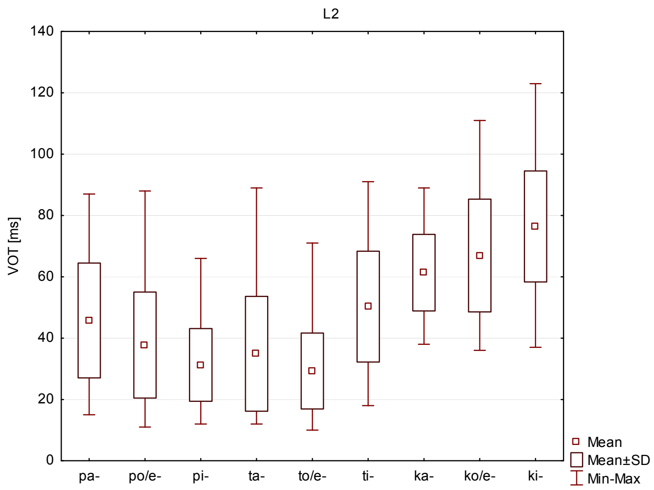


Figure 72. Box plot of mean VOT values for various vocalic contexts in L2 German

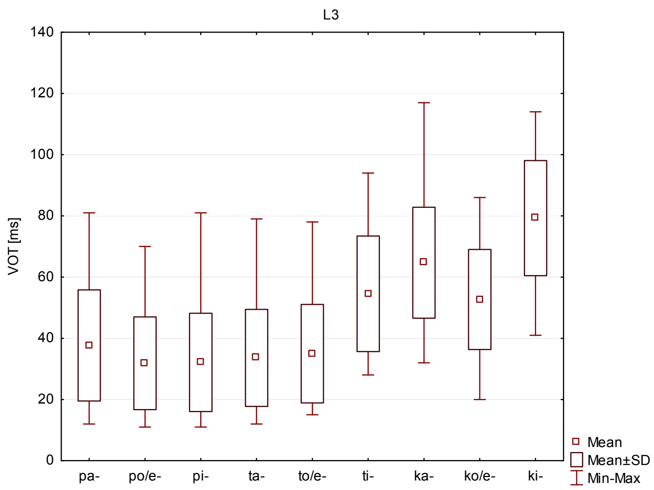


Figure 73. Box plot of mean VOT values for various vocalic contexts in L3 English

5.4.3.10. Analysis of variance

A two-factor ANOVA between the languages (L1, L2, L3) and the VOT durations of the voiceless plosive sounds /p, t, k/ was performed as part of the analysis of variance.

Table 117. Two-factor analysis of variance for /p, t, k/ and for L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	96.56 (0.05;2;1746)	0.000000*
plosive (/p/. /t/. /k/)	531.43 (0.05;2;1746)	0.000000*
Language*plosive	5.01 (0.05;4;1746)	0.000515*

*p<0.05

The results indicate that the differences in the VOT values within the factors of the language (F (2; 1746) =96.56, p<.05) and the plosives (F (2; 1746) =531.43, p<.05) were found to be significant. Moreover, the interaction between the languages and the plosives on the VOT values was shown to be significant (F (4; 1746) =5.01, p<.05). It can be interpreted that there are significant differences between the VOT values for at least two languages, and for at least two plosives. The existing interaction between the languages and the plosives depends on the type of language (L1, L2, L3), as presented in Table 117 and Figure 74.

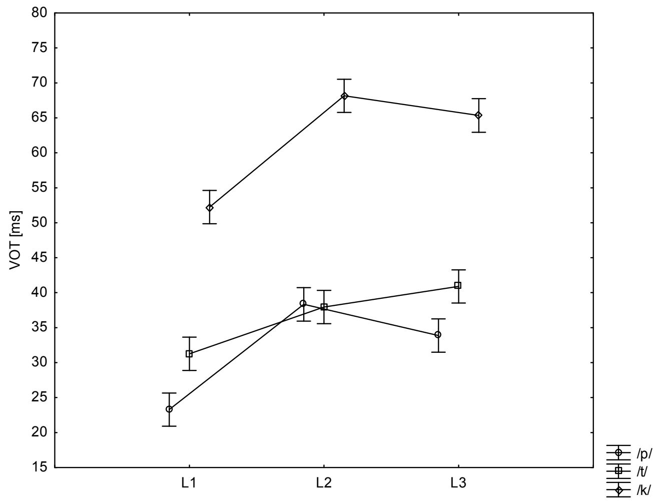


Figure 74. Interaction between the language and plosive factors

A further analysis was aimed at investigating the interaction of the vowel context and the language on the observed VOT durations in Group C. To this end, a two-factor analysis ANOVA was performed for the factors of the language (L1, L2, L3) and the context of the vowel following the voiceless plosives in the stressed onset positions in the target words (i.e. low _/a/, high _/i/, and mid _/e, o/). In accordance with the universal tendencies, the context of the high vowels (e.g. /i/) should generate longer VOT values in the preceding plosives than the context of the low vowels (e.g. /a/). The results of the analysis indicate that there are significant differences in the VOT values within the factor of the languages (F (2; 1746) =63.9, p<.05) and the vowel context (F (2; 1746) =49.2, p<.05). It follows that significant differences between the VOT values were found for at least two languages, and for at least two vowel contexts. Furthermore, a significant interaction was found between the two factors (F (4; 1746) =3.37, p<.05), which depends on the type of the language (see Table 118 and Figure 75).

Table 118. Two-factor analysis of variance for vowel context /-a-/, /-o/e-/, /-i-/ and L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	63.891 (0.05;2;1746)	0.000000*
Vowel (/-a-/. /-o/e-/. /-i-/)	49.188 (0.05;2;1746)	0.000000*
Language*Vowel	3.367 (0.05;4;1746)	0.009382*

*p<0.05

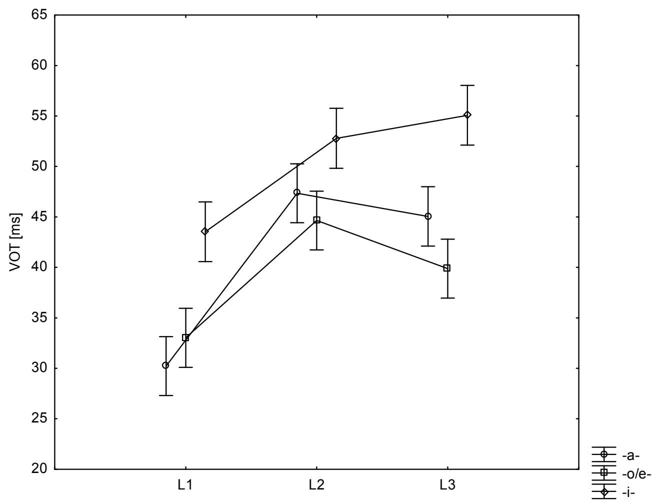


Figure 75. Interaction between the language and vowel context factors

Different patterns of interaction can be observed in the respective languages with only L1 Polish following closely the universal patterns, i.e. the longest VOT values in the high vowel context /i/, medium for the mid vowels /e/, /o/, and the shortest for the low vowel context /a/. In L2 German and L3 English the universal tendencies were partially observed, with the high vowel context generating, on average, the longest VOT values, whereas the mid and low vowel contexts yielded slightly different VOT duration patterns than the expected ones, i.e. the mid vowel contrasts /-o/e-/ were found to be shorter than the low vowel context /-a-/.

With the view of exploring the effects of language proficiency on the generated VOT values in all the languages, the participants in Group C were divided into two subgroups, including the 1EG lower proficiency group and the 2/3/4EG higher proficiency group. The results of the two-factor ANOVA for the proficiency groups and for the L1, L2, L3 demonstrated that the differences in the VOT values within the factors of language ($F (2; 1746)=59.7$, $p<.05$) were significant; however, there were no significant differences between the mean VOT values with respect to the proficiency group ($F (1; 1746) =0.07$, $p>.05$). Further, there was a significant interaction between the factors of the language and the group on the mean VOT values ($F (2; 1746) =8.36$, $p<.05$) and it depended on the type of language (see Table 119 and Figure 76).

Table 119. Two-factor analysis of variance for proficiency groups and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	59.764 (0.05;2;1749)	0.000000*
Group (1EG, 2/3/4EG)	0.067 (0.05;1;1749)	0.796297
Language*Group	8.360 (0.05;2;1749)	0.000244*

*p<0.05

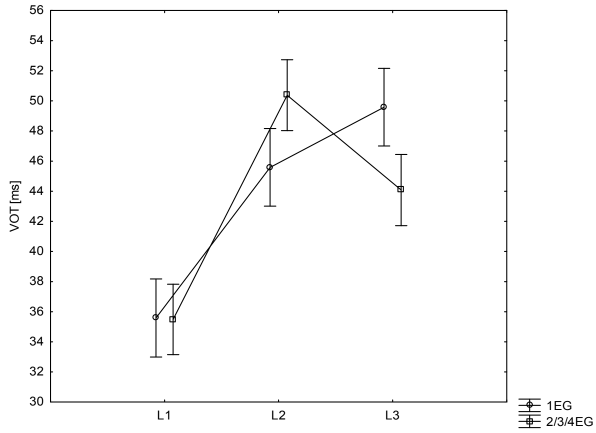


Figure 76. Interaction between the language and proficiency group factors

The findings of the two-factor ANOVA analysis of variance for the proficiency groups and the place of articulation PoA /p, t, k/ separately for all the languages (L1, L2, L3) are as follows. For L1 Polish, there were statistically significant differences between the mean VOT values for at least two places of articulation (F (2; 581) =242.8, p<.05); there were no significant differences in VOT for the proficiency group; there was a significant interaction between the PoA and the group(F (2; 581) =3.6, p<.05).For L2 German and L3 English, there were statistically significant differences between the mean VOT values for at least two places of articulation, i.e. for L2(F (2; 579) =183, p<.05) and for L3 (F (2; 577) =146, p<.05); there were statistically significant differences between the mean VOT values for the proficiency group, i.e. for L2 (F (2; 579) =10.5, p<.05) and for L3 (F (2; 577) =11.6, p<.05); there were no significant interactions between the PoA and the groups (see Table 120 and Figure 77).

Table 120. Two-factor analysis of variance for proficiency groups and place of articulation PoA /p, t, k/ and for languages L1, L2, L3

Factor	L1		L2		L3	
	F	p	F	p	F	p
PoA (/p/, /t/, /k/)	242.852 (0.05;2;581)	0.000000*	183.156 (0.05;2;579)	0.000000*	146.358 (0.05;2;577)	0.000000*
Group (1EG, 2/3/4EG)	0.004 (0.05;1;581)	0.949774	10.462 (0.05;1;579)	0.001288*	11.645 (0.05;1;577)	0.000689*
PoA*Group	3.632 (0.05;2;581)	0.027061*	0.526 (0.05;2;579)	0.590965	0.848 (0.05;2;577)	0.428660

*p<0.05

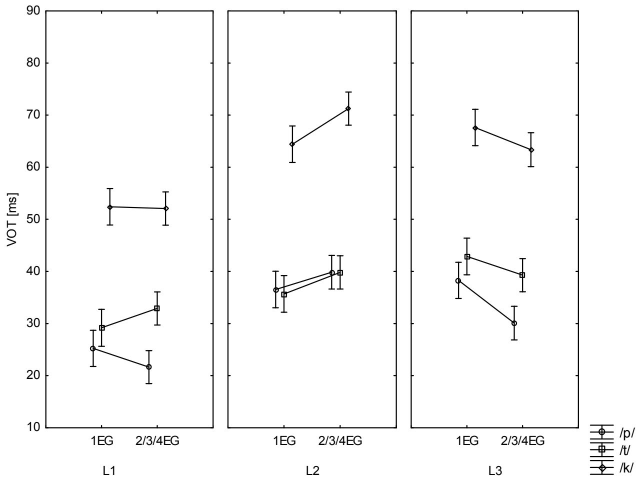


Figure 77. Interaction of PoA (/p/, /t/, /k/) and Group (1EG, 2/3/4EG) for languages L1, L2 and L3

5.4.3.11. Multiple regression analysis

Multiple regression analyses were performed to estimate the relationship between the dependent and independent variables in Group C.

The first regression analysis investigated the relationship between the VOT values as the dependent variable and a number of independent variables, including the language (L1, L2, L3), the place of articulation PoA (/p/, /t/, /k/), the vowel context and the proficiency group. The assumed levels of statistical significance pointed to a statistically significant interdependence between the VOT values and the following variables of the language, the place of articulation and the vowel context, whereas the influence of the group variable was not found to be significant. The R squared result indicated that 46.7% of the variance is accounted for by the independent variables (see Table 121).

Further, a standardized coefficient, i.e. the beta coefficient was calculated to establish which of the independent variables had a greater effect on the dependent variable in the multiple regression analysis. The beta coefficient values suggest that PoA 1, i.e. /p/ (b=0.4), Language 1, i.e. L1 (b=0.29) and PoA 2, i.e. /t, k/ (b=0.25) have the biggest influence on the VOT values, whereas the impact of other languages (L2, L3) as well as the vowel context (Vowel 2 -i- and Vowel 1 -a-, -o/e-) is relatively smaller (see Table 122).

Table 121. Multiple regression analysis 1

Variable	R	R^2	F	p
VOT	0.68	0.467	218.3 (0.05;7;1747)	0.000000*

*p<0.05

Table 122. Results of multiple regression 1 for variables

Effect	t	p	Beta
Target word	111.67	0.000000*	–
Language 1	-14.42	0.000000*	-0.291
Language 2	8.55	0.000000*	0.172
Group	0.31	0.758594	0.005
PoA 1	-21.30	0.000000*	-0.429
PoA 2	-12.33	0.000000*	-0.248
Vowel 1	-4.78	0.000002*	-0.096
Vowel 2	-8.03	0.000000*	-0.161

*p<0.05

Another multiple regression was performed for the L3 VOT values as the dependent variable and the L1 VOT and the L2 VOT as the independent variables. The analysis pointed to a statistically significant interdependence among the variables under investigation. The regression model is explained in 39.1% by the independent variables L1 VOT and L2 VOT (R^2=0.391).The bigger impact on the dependent variable L3 VOT is exerted by the L2 VOT values (b=0.42) than the L1 VOT (b=0.29) (see Tables 123 and 124).

Table 123. Multiple regression analysis 2

Variable	R	R^2	F	p
L3 VOT	0.63	0.391	181.6 (0.05;2;565)	0.000000*

*p<0.05

Table 124. Results of multiple regression 2 for variables

Effect	t	p	Beta
Target word	6.36	0.000000*	–

L1 VOT	7.62	0.000000*	0.292
L2 VOT	11.01	0.000000*	0.422

*p<0.05

Finally, a regression analysis was calculated for the L3 VOT values as the dependent variable and other independent variables including the L1 VOT, the L2 VOT, the place of articulation /p, t, k/, the vowel context and the proficiency group. The regression model is explained in 49% by the independent variables. The assumed level of testing probability points to a statistically significant impact of such independent variables as the L2 VOT, the PoA /p, k/, the vowel contexts /-o/e-/, /-i-/ and the group, whereas the interdependence is insignificant for the remaining variables. The greatest impact is exerted by the L2 VOT (b=0.33) and the place of articulation /p, k/ (b=0.3), whereas it is smaller in the case of the vowel context /-o/e-/, /-i-/ (b=0.19) and the group (b=0.15) (see Tables 125 and 126).

Table 125. Multiple regression analysis 3

Factor	R	R^2	F	p
L3 VOT	0.70	0.490	76.8 (0.05;7;560)	0.000000*

*p<0.05

Table 126. Results of multiple regression 3 for variables

Effect	t	p	Beta
Target word	10.31	0.000000*	–
VOT L1	1.15	0.249547	0.053
VOT L2	8.19	0.000000*	0.328
Group	4.94	0.000001*	0.151
PoA 1	-7.01	0.000000*	-0.309
PoA 2	-1.73	0.084636	-0.066
Vowel 1	-1.04	0.298660	-0.038
Vowel 2	-5.57	0.000000*	-0.196

*p<0.05

5.4.4. Results for Group D

The results of the VOT analyses for Group D with L1 Polish, L2 French and L3 English will be presented in the following sections.

5.4.4.1. Mean VOT values for L1, L2 and L3

The Shapiro–Wilk test of normality was administered to verify the null hypothesis for this test that the VOT measurements data are normally distributed ($p>\alpha=0.05$). Since the chosen alpha level was 0.05 and the p-values for the VOT variable in several categories were less than 0.05, then the null hypothesis that the data are normally distributed had to be rejected. Consequently, non-parametric tests were applied for the further statistical analyses of the data in Group D.

Tables 127-129 present the mean results of the VOT measurements of the voiceless plosives /p/, /t/, /k/ of the target words read in the carrier phrases in the participants' L1 Polish, L2 French and L3 English in Group D.

Table 127. Mean VOT values for target words in L1 Polish

Words in L1	N	Mean	SD	Median
pan	29	20.4	11.3	17.0
para	29	19.8	10.8	16.0
pech	29	16.9	6.9	15.0
pole	29	24.3	10.4	21.0
pik	29	24.3	10.6	24.0
piwo	29	31.9	13.0	34.0
tak	29	21.1	7.3	20.0
tango	29	23.7	8.7	20.0
tor	29	26.2	10.0	24.0
testy	29	22.7	6.6	23.0
tik	29	39.2	13.0	37.0
tiry	29	43.2	17.3	42.0
kat	29	45.3	11.5	45.0
kanon	29	43.7	11.4	41.0
kot	29	45.7	10.6	45.0
kelner	29	41.1	11.7	40.0
kit	29	57.7	12.7	57.0
kino	29	59.4	14.0	62.0

Table 128. Mean VOT values for target words in L2 French

Words in L2	N	Mean	SD	Median
pas	28	18.7	11.6	17.0
part	28	16.5	7.6	15.0
poche	28	22.1	8.9	19.0
perdre	27	17.9	10.0	15.0

pic	28	31.5	12.2	26.0
pile	28	47.5	20.6	41.5
taxe	28	22.8	7.6	22.0
table	28	25.6	8.5	24.0
terre	28	28.6	10.5	26.0
torse	28	29.9	10.6	30.0
tic	27	53.8	19.3	51.0
titre	27	47.8	17.4	44.0
calme	28	51.4	18.7	51.0
carte	28	42.3	12.7	42.0
coq	28	40.7	10.9	41.0
quelle	28	44.7	14.3	41.0
qui	28	79.7	21.4	79.5
quiche	23	56.6	11.0	54.0

Table 129. Mean VOT values for target words in L3 English

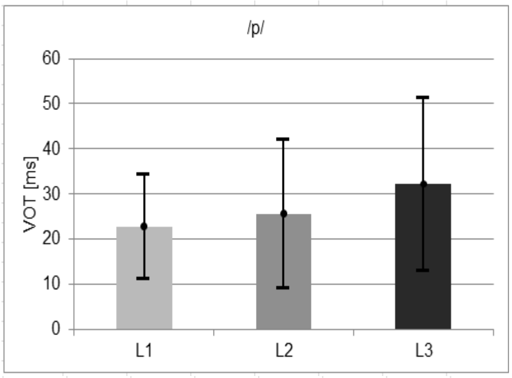
Words in L3	N	Mean	SD	Median
pan	28	36.0	26.0	29.5
party	28	30.3	16.5	28.0
pot	28	36.1	20.2	30.0
pencil	28	25.9	14.4	24.5
peace	28	32.9	18.5	28.0
Peter	28	32.8	17.2	30.5
task	28	37.4	22.6	33.0
taxi	28	26.7	13.7	23.0
test	28	37.2	20.2	33.0
toffee	28	37.4	22.9	27.5
team	28	55.9	24.3	52.0
teacher	28	50.1	16.9	50.0
cat	28	75.0	30.3	74.0
carpark	28	61.4	24.6	61.0
kept	28	48.0	21.7	43.0
coffee	28	53.4	19.4	52.0
keen	28	86.0	18.7	86.5
keeper	28	68.7	22.9	64.0

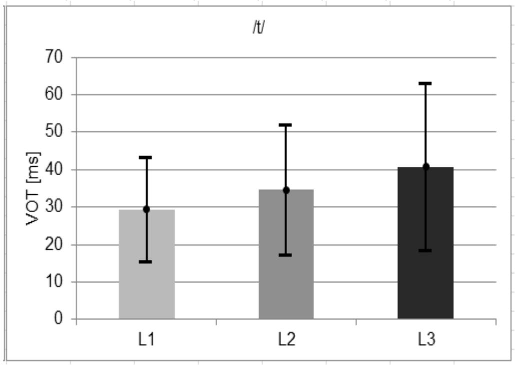
The acoustic measurements of the mean voice onset time for the voiceless plosives /p/, /t/, /k/ in the stress onset positions were performed in the participants' L1 Polish, L2 French and L3 English. The results are presented in Table 130 and Figures 78-80.

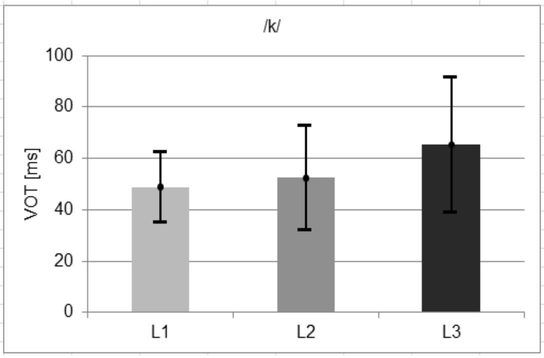
The durational measurements demonstrated that the multilingual participants produced the voiceless plosives in the stress onset positions with the mean voice onset time values that were comparable in L1 Polish (/p/=23 ms, /t/=29 ms, /k/=49 ms) and L2 French (/p/=26 ms, /t/=35 ms, /k/=52 ms), whereas they had a somewhat longer-lag in L3 English (/p/=32 ms, /t/=41 ms, /k/=65 ms).

Table 130. Mean VOT values for L1, L2 and L3.

VOT	Language	N	Mean	SD	Median
/p/	L1	174	22.9	11.5	20.0
	L2	167	25.7	16.5	21.0
	L3	168	32.3	19.2	28.0
/t/	L1	174	29.4	13.9	26.5
	L2	166	34.6	17.3	31.0
	L3	168	40.8	22.3	36.0
/k/	L1	174	48.8	13.8	46.0
	L2	163	52.4	20.4	49.0
	L3	168	65.4	26.3	62.0







Figures 78-80. VOT measurements for /p, t, k/ in L1 Polish, L2 French, L3 English

5.4.4.2. Cross-language comparison of VOT means

A series of statistical tests was run with the view of investigating the language effect on the VOT durations in Group D. A non-parametric Kruskal-Wallis ANOVA was conducted in order to compare the mean VOT values for /p/, /t/, /k/ across the participants' languages (L1, L2, L3) with the assumed alpha level being 0.05.

The results of the ANOVA pointed to significantly different values for all the initial voiceless plosives between L1 Polish and L3 English as well as between L2 French and L3 English ($p<.05$). However, the differences between the mean VOT values for two out of three plosives in L1 Polish and L2 French were not found to be statistically significant (see Table 131). A detailed analysis of the cross-language comparison of the VOT means is presented in the following section.

Table 131. Mean VOT comparison for /p t k/ in L1 Polish, L2 French, L3 English

Kruskal-Wallis Test	p for multiple comparisons		
	L1 vs L2	L1 vs L3	L2 vs L3
/p/	0.912010	0.000002*	0.000238*
/t/	0.014929*	0.000000*	0.026762*
/k/	0.837185	0.000000*	0.000001*

*p<0.05

With respect to the cross-linguistic measurements for the bilabial plosive /p/, the p values for the multiple comparisons pointed to statistically significant differences in the mean VOT between the following language pairings: between the L1 and the L3 (i.e. the mean VOT values in L1 Polish were lower than in L3 English); between the L2 and the L3 (i.e. the mean VOT values in L2 French were lower than in L3 English). The difference between the mean values of /p/ in L1 Polish and L2 French was not found to be significant.

As far as the mean VOT values for /t/ are concerned, the multiple comparisons demonstrated statistically significant differences between all the languages involved, i.e. between the L1 and the L2 (i.e. the mean VOT values in L1 Polish were lower than in L2 French); between the L1 and the L3 (i.e. the mean VOT values in L1 Polish were lower than in L3 English); between the L2 and the L3 (i.e. the mean VOT values in L2 French were lower than in L3 English).

In the case of the mean VOT durations for the velar plosive /k/, similar cross-linguistic patterns were observed as for /t/ measurements. The performed multiple comparisons also demonstrated statistically significant differences between the following language pairings: between the L1 and the L3 (i.e. the mean VOT values in L1 Polish were lower than in L3 English); between the L2 and the L3 (i.e. the mean VOT values in L2 French were lower than in L3 English). However, the difference between the mean VOT values for /k/ in L1 Polish and L2 French did not prove to be significant.

Summing up, the performed tests yielded the cross-linguistic differences that proved to be significant between the third language (i.e., English) and native Polish, as well as between both non-native languages (i.e. L3 English and L2 French). Further, the pairwise comparison of means between L1 Polish and L2 French demonstrated a significant difference only for /t/ ($p<0.05$), whereas the differences for the bilabial and velar plosives /p/ and /k/ were found to be non-significant.

The following box plots (Figures 81-83) illustrate the tendencies observed in the VOT patterns in the respective languages separately for the stressed onset plosives /p/, /t/ and /k/. Particularly noteworthy is that the distributions in L1 Polish and L2 French show similar trends with respect to the VOT means and the standard deviation, whereas the mean values for L3 English remain significantly higher and the minimum-maximum range is much more extreme.

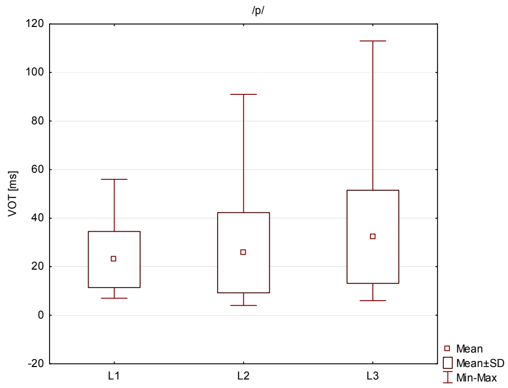


Figure 81. Box plot of mean VOT values for /p/ in L1 Polish, L2 French and L3 English

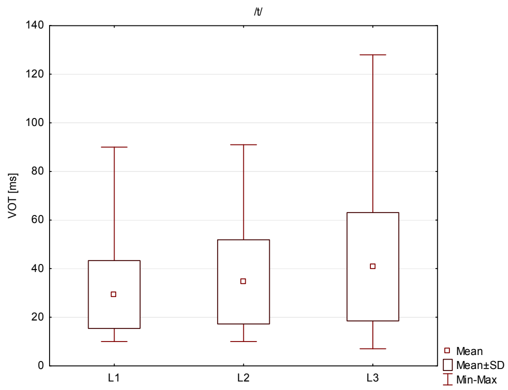


Figure 82. Box plot of mean VOT values for /t/ in L1 Polish, L2 French and L3 English

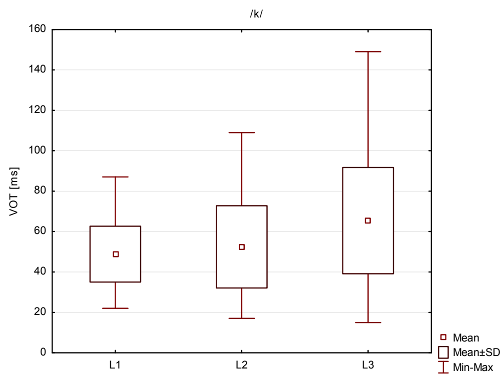


Figure 83. Box plot of mean VOT values for /k/ in L1 Polish, L2 French and L3 English

Recapitulating, the language effect was observed to occur between the third language and the remaining native and non-native phonological systems in Group D. The mean voice onset time values in L3 English were significantly higher than the respective values in L1 Polish and L2 French, which displayed rather similar patterns of distribution.

5.4.4.3. Proficiency group effect

The language proficiency group effect on the VOT acquisition was examined by means of the non-parametric Mann-Whitney U tests, which were performed for the mean VOT values for /p/, /t/, /k/ in L1, L2 and L3 for two proficiency subgroups in Group D, i.e. the 1E lower proficiency group (N=9) vs. the 2/3E higher proficiency group (N=19) (see Table 132).

Table 132. Results of Mann-Whitney test between two groups 1E vs. 2/3E for /p, t, k/ in L1, L2 and L3

Sounds	Language	Z corr.	p
/p/	L1	0.33	0.738601
	L2	0.12	0.900544
	L3	-1.06	0.287512
/t/	L1	0.12	0.908014
	L2	0.68	0.496938
	L3	-1.54	0.124193
/k/	L1	-0.15	0.877150
	L2	0.91	0.362814
	L3	-1.28	0.200906

*p<0.05

The results of the analysis failed to demonstrate any significant proficiency group effects with respect to differences between the mean VOT for /p, t, k/ in L1 Polish, L2 French and L3 English in Group D.

5.4.4.4. Cross-linguistic correlations between VOT values

The Pearson product-moment correlation analysis was performed in order to investigate the dependence between two variables, i.e. the observed VOT values in L3 English and those of the native L1 Polish as well as between L3 English and L2 French (the assume alpha level was α=0.05). The calculated coefficients pointed to prevailingly positive weak to moderate correlations between the mean VOT values in Group D.

For the voiceless bilabial plosive /p/, there were statistically significant correlations for both pairs of variables. The results pointed to a weak positive correlation (R=0.18) between L1 Polish and L3 English as well as a weak correlation (R=0.19) between both non-native languages in the mean VOT values for /p/ see Table 133.

Table 133. Pearson's correlation for /p/

Pair of variables for /p/	N	R	t	p
VOT L1 & VOT L3	167	0.18	2.35	0.019988*
VOT L2 & VOT L3	167	0.19	2.51	0.013053*

*p<0.05

As the correlations between the VOT durations for both language pairing are significant in the case of all plosives and they are rather similar in magnitude, only one selected scatterplot will be presented representing the pair of variables with a higher correlation coefficient, i.e., L2 French and L3 English (see Figure 84).

In the case of the alveolar plosive /t/, the obtained Pearson's correlations for both pairs of variables were in the weak range; with the coefficient for L1 Polish and L3 English being R=0.32, and for L2 French and L3 English R=0.36 (see Table 134).

Table 134. Pearson's correlation for /t/

Pair of variables for /t/	N	R	t	p
VOT L1 & VOT L3	166	0.32	4.31	0.000028*
VOT L2 & VOT L3	166	0.36	4.86	0.000003*

*p<0.05

As far as the Pearson's correlations for VOT in the velar plosive /k/ are concerned, a weak positive correlation was found to hold between L1 Polish and L3 English (R=0.2). In the case of the second pair of variables, i.e. L2 French and L3 English, the coefficient was significantly higher and pointed to a moderate correlation between the non-native languages (R=0.45) (see Table 135, Figure 84).

Table 135. Pearson's correlation for /k/

Pair of variables for /k/	N	R	t	p
VOT L1 & VOT L3	163	0.20	2.54	0.012123*
VOT L2 & VOT L3	163	0.45	6.35	0.000000*

*p<0.05

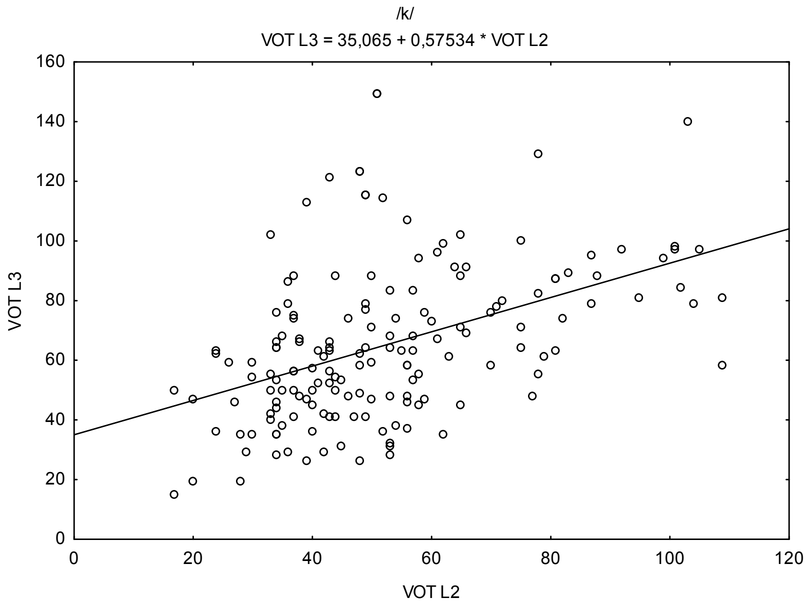


Figure 84. Scatterplot of VOT for /k/ between L2 French and L3 English

Since the covariance between the respective variables remains rather weak, with a tendency for slightly stronger correlations between the L2 French and L3 English VOT values, compared to the dependency between L1 Polish and L3 English. Therefore, it cannot be concluded with certainty whether the L3 VOT values in Group D correlated more with the native values (L1) or those of the previously acquired foreign language (L3).

5.4.4.5. Comparison to VOT reference values

Further statistical analyses were performed including one-sample t-tests and a non-parametric Wilcoxon signed-rank matched pairs test in order to compare the acoustic measurements obtained in Group D for /p, t, k/ in L1 Polish, L2 French and L3 English to the monolingual reference values quoted in the literature, i.e. Keating et al. (1981) for Polish; Caramazza et al. (1973) for French and Lisker and Abramson (1964) for English.

Table 136. Comparison to VOT reference values in Polish, French and English ([1]Keating et al., 1981; [2]Caramazza et al. 1973; [3]Lisker and Abramson, 1964)

Language	Parameter	VOT		
		/p/	/t/	/k/
Polish	Ref. VOT[1]	22	28	52
	L1 Mean	22.9	29.4	48.8
	SD	11.5	13.9	13.8
	p	0.291950	0.201095	0.002898*
French	Ref. VOT[2]	18	23	32
	L2 Mean	25.7	34.6	52.4
	SD	16.5	17.3	20.4
	p	0.000000*	0.000000*	0.000000*
English	Ref. VOT[3]	59	67	84
	L3 Mean	32.3	40.8	65.4
	SD	19.2	22.3	26.3
	p	0.000000*	0.000000*	0.000000*

*$p<0.05$

The findings show that the VOT durations obtained in the study for L2 French and L3 English differ significantly from the monolingual reference values as reported in the literature, whereas the VOT values for L1 Polish correspond to the reported norms for Polish (see Table 136). In fact, the VOT values for the voiceless stops in L1 Polish were very close to the monolingual norms from the literature (cf. Keating et al., 1981), i.e. nearly identical in the case of /p/ and /t/ and slightly lower for /k/ (49 vs. 52 ms), although still within the accepted 5 ms range.

Both in L2 French and L3 English the observed values differed from the monolingual reference values; however, the directionality of the discrepancy differed. In the case of L2 French, the VOT values generated in the study exhibited longer mean duration of voice onset time that the literature reference values (Caramazza et al. 1973), i.e. 26 vs. 18 ms for /p/, 35 vs. 23 ms for /t/, 52 vs. 32 ms for /k/. In L3 English, on the other hand, the onset voiceless plosives were realized with a shorter lag than the monolingual norms (Lisker and Abramson, 1964), with 32 vs. 59 ms for /p/, 41 vs. 67 ms for /t/, and 65 vs. 84 ms for /k/.

It is interesting to note that the mean VOT measurements in both non-native languages exhibited intermediate values and did not approximate fully the respective native target norms. Particularly noteworthy is that the L2 French VOT values were longer than the L1 Polish as well as the target native French durations, so there was no facilitation by L1 transfer in this respect, as could have been expected. An explanation for the longer lag found in the L2 French values could be a potential influence of the L3 English VOT values. What should be borne in mind is the questionable status of English as the third language in the case of Group D. Although French has been the dominant foreign language for the participants at the moment of testing, in several cases English was acquired earlier but lost its prevalent status in the course of time.

The VOT values in L3 English were considerably higher than in L2 French, yet did not reach the native English level. The multilingual participants in Group D seem to be aware of the differences in the phonetic realisation of the laryngeal contrasts in different languages and try to keep their language systems apart, although they do not necessarily approximate the target reference values.

The native system of the participants does not appear to be affected by other non-native languages as the VOT durations for L1 Polish quite closely resembled those of the Polish monolingual norms.

5.4.4.6. Comparison to control groups

With a view to verifying further the reference values found in the literature, the VOT measurements generated in Group D were compared to the corresponding control groups involved in the study. The French control group consisted of 8 native French speakers (i.e. lecturers at Adam Mickiewicz University in Poznań), who were recorded reading the word list in the carrier phrases for French. The English control group involved 17 native English speakers (i.e. lecturers at the University of Essex and Adam Mickiewicz University in Poznań), who were recorded reading the list for English.

The generated control mean VOT values were /p/=29, /t/=36, /k/=56 ms for French and /p/=61, /t/=75, /k/=79 ms for English. In the case of the English control group, the VOT values were in the range of the reference values from the literature. However, the French control group measurements exceeded by more than 10 ms the monolingual reference values for French from the literature (see Table 137).

In order to compare the observed mean VOT durations for /p, t, k/ in the participants' L2 French and L3 English to the respective control groups, one-sample t-tests and a non-parametric Wilcoxon signed-rank test for matched pairs were performed ($\alpha=0.05$).

Table 137. Comparison of mean VOT for /p, t, k/ in L2 French and L3 English to Control Groups values

Language	Parameter	VOT		
		/p/	/t/	/k/
French	French Controls	29	35	56
	L2 Mean	25.7	34.6	52.4
	SD	16.5	17.3	20.4
	p	0.000010*	0.068707	0.001850*
English	English Controls	61	75	79
	L3 Mean	32.3	40.8	65.4
	SD	19.2	22.3	26.3
	p	0.000000*	0.000000*	0.000000*

*$p<0.05$

As far as L2 French is concerned, the statistical test pointed to a significant difference between the VOT durations of the participants and the French native control group in the case of /p/ and /k/. However, when analyzing the results more closely, the duration differences were still within the acceptable 5 ms range, thus, we can assume that the participants' VOT values approximated those of the native French controls, the former being slightly lower (26 vs. 29 ms for /p/, 35 vs. 35 ms for /t/, 52 vs. 56 ms for /k/).

In the case of L3 English, statistically significant differences were found for the comparison of all the VOT durations with the multilingual participants implementing the voiceless plosives with considerably shorter lag values than the native English controls. As stated earlier, the English controls' VOT values were well within the range reported in the literature, thus confirming the validity of the selected reference norms.

Recapitulating, when compared to the native control groups, the participants in Group D approximated the target VOT patterns in their L2 French, whereas they failed to do so in L3 English.

5.4.4.7. VOT goodness of fit

In order to evaluate the participants' approximation to the target control values in Group D, the same procedure was followed as in Groups A, B and C. The calculation of the measure of the nativeness effect or the VOT goodness of fit, based on the degree of approximation or deviation from the control VOT values, is explained in section 2.1.7. Consequently, the participants' measurements were assigned to particular categories, the higher the number the more they differed from the control baseline durations, with '+ values' indicating a longer lag and '- values' shorter VOT durations. Table 138 shows the distribution of percentage scores across the approximation threshold levels for L2 French and L3 English in comparison to the respective control VOT durations.

Table 138. Level of approximation of L2 and L3 VOT values to control VOT (in % of participants)

Approximation level	L2 French to Control VOT			L3 English to Control VOT		
	/p/	/t/	/k/	/p/	/t/	/k/
-9	0	0	0	0.6	0.6	0.6
-8	1.2	0	0	6.0	1.8	1.2
-7	3.0	1.2	0	17.9	13.7	5.4
-6	9.0	3.0	2.5	16.1	23.8	7.1
-5	11.4	6.0	3.1	17.9	13.1	11.9
-4	12.6	7.8	6.7	8.9	11.3	13.1
-3	11.4	17.5	17.8	10.7	14.9	14.9
-2	5.4	6.0	12.3	7.1	7.7	11.3
-1	12.6	12.7	12.9	4.2	3.6	13.7
0	12.0	13.3	19.6	4.2	3.6	7.1
1	3.6	4.2	6.1	1.2	1.8	6.0
2	1.2	8.4	2.5	2.4	0.6	0.6
3	0.6	2.4	4.3	0.6	1.2	2.4
4	1.8	4.2	3.7	0.6	0.6	1.8
5	4.2	2.4	1.8	0.6	0.6	1.2
6	1.2	3.0	1.2	0.6	0.6	0.6
7	0.6	0.6	0.6	0.6	0.6	1.2
8	0.6	0.6	3.7	0	0	0
9	0.6	0	1.2	0	0	0
10	1.2	1.8	0	0	0	0
11	1.8	1.2	0	0	0	0
12	0.6	0	0	0	0	0
13	1.8	1.8	0	0	0	0
14	1.2	1.2	0	0	0	0
15	0.6	0.6	0	0	0	0

Table 139 presents the mean results of the approximation level analysis and points to diverse scores for L2 French (M=-0.5) and L3 English (M=-3.3). It appears that the L1 nativeness effect, based on the VOT approximation to the respective control values, was more visible in the case of L2 French, in which the VOT measurements were in a close range of the control native values, whereas in the case of L3 English, the VOT durations deviated more significantly from the baseline values.

Table 139. Approximation level mean results

Language	N	Mean	SD	Median	Minimum	Maximum
L2	496	-0.5	4.49	-1.0	-8.0	21.0
L3	504	-3.3	3.17	-4.0	-9.0	8.0

5.4.4.8. Individual variation

The VOT measurements were also analyzed with respect to the individual variation in the generated VOT values for /p/ /t/ /k/ in Group D. The following figures illustrate the individual variation separately for the L1, the L2 and the L3 against the selected reference VOT values for particular languages.

As can be seen in Figures 85-87, the participants followed the universal VOT patterns in all their respective languages (L1, L2, L3) with the bilabial plosives yielding the shortest VOT values, and the velar – the longest.

The greatest variability in the VOT distribution can be observed in L3 English followed by L2 French. In general, the L3 values are undershoots of the reference English VOT measurements, with the most extreme departures from the target represented by such individuals as 1E_MG, 1E_JO, 2E_JT, 2E_MP, 3E_ND. Only in a few cases, the L3 performance approximates closely the native English VOT durations (1E_MPO, 1E_ABA), and there is one participant with the overshoot values for all the plosives, i.e. 2E_MK.

There was quite a considerable variability exhibited also in L2 French and no clear correspondence was noticeable between the participants' laryngeal contrasts duration in their respective non-native languages. As expected, the VOT durations for L1 Polish exhibited less interspeaker variation and a fairly close correspondence to the VOT reference values for Polish.

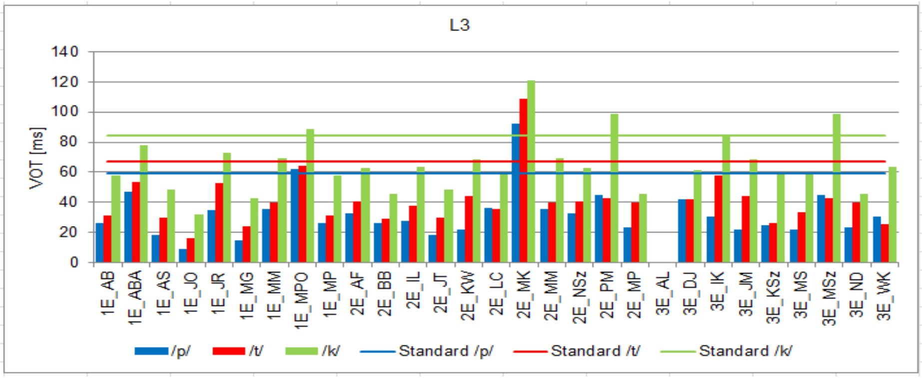


Figure 85. Individual variation in L3 English VOT for /p/ /t/ /k/ against the reference VOT values

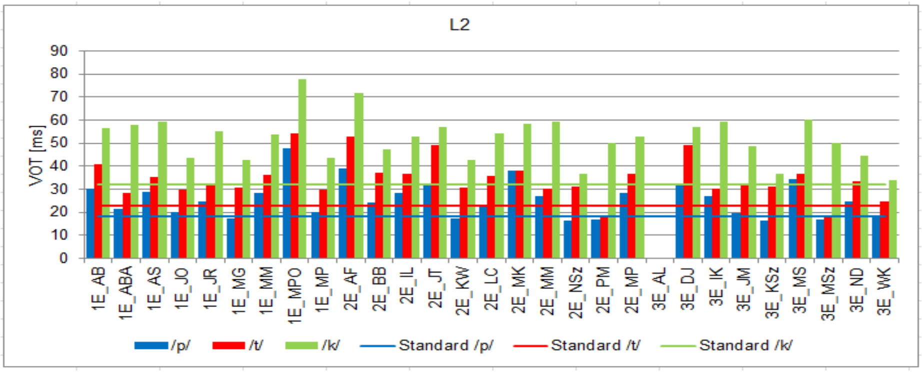


Figure 86. Individual variation in L2 French VOT for /p/ /t/ /k/ against the reference VOT values

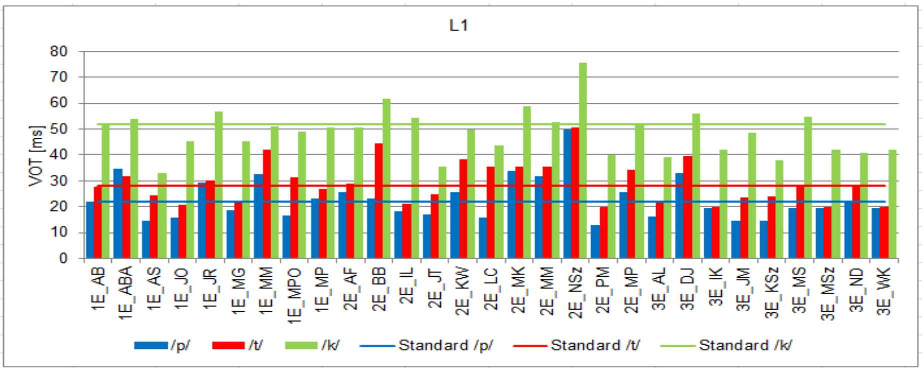


Figure 87. Individual variation in L1 Polish VOT for /p/ /t/ /k/ against the reference VOT values

5.4.4.9. Vocalic context effects

A non-parametric Kruskal-Wallis test was used to investigate differences in the mean VOT with respect to the vocalic context in each of the participants' languages in Group D. The assumed alpha level was $\alpha=0.05$.

In the case of the bilabial voiceless plosive /p/, the results of the performed Kruskal-Wallis test point to the statistically significant differences in the mean VOT values between /pa-, po/e-, pi-/ contrasts for L1 Polish and L2 French, i.e. with the low vowel /a/ and the mid vowels /o, e/ preceding the initial plosives generating shorter VOT durations compared to the high vowel /i/ contexts. Such observed patterns are in line with the universal tendencies in the vocalic context effects on the VOT durations. No statistically significant differences between the high vs. mid vs. low vowel contexts were found in L3 English (see Table 140).

Table 140. Comparison of mean VOT between /pa-, po/e-, pi-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	pa- vs po/e-	pa- vs pi-	po/e- vs pi-
L1	0.0002*	1.000000	0.000235*	0.004761*
L2	0.0000*	0.516176	0.000000*	0.000000*
L3	0.8028	1.000000	1.000000	1.000000

*p<0.05

The analysis of the vocalic context effects for the alveolar plosive /t/ demonstrated statistically significant differences in the mean VOT values for all the languages, i.e. the L1, L2, L3. As in the previous case, the low vowel /a/ and the mid vowels /o, e/ following the initial plosives generate shorter VOT durations compared to the high vowel /i/ contexts; however, there was no significant difference in the VOT durations between the low and the mid vowels contexts (see Table 141).

Table 141. Comparison of mean VOT between /ta-, to/e-, ti-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ta- vs to/e-	ta- vs ti-	to/e- vs ti-
L1	0.000000*	0.746842	0.000000*	0.000000*
L2	0.000000*	0.101456	0.000000*	0.000000*
L3	0.000000*	0.202654	0.000000*	0.000015*

*p<0.05

The results of the Kruskal-Wallis test for the multiple comparisons between the vocalic contexts for the velar plosive /k/ showed statistically significant differences in the mean VOT measures for all the languages involved. For L1 Polish and L2 French, consistently with the previous results, the high /i/ vowel context resulted in longer VOT durations of the preceding consonant compared to both low and mid vowel contexts. A slightly different pattern was observed in the case of L3 English, where significant differences in the VOT durations were found for the comparisons between the high and mid vowels as well as the low and mid vowel contexts (see Table 142).

Table 142. Comparison of mean VOT between /ka-, ko/e-, ki-/ contexts for L1, L2 and L3

Language	Kruskal-Wallis Test	p for multiple comparisons		
	p	ka- vs ko/e-	ka- vs ki-	ko/e- vs ki-
L1	0.000000*	1.000000	0.000000*	0.000000*
L2	0.000000*	0.686689	0.000000*	0.000000*
L3	0.000000*	0.000673*	0.054643	0.000000*

*p<0.05

On the whole, the vocalic context effects observed in the data follow, to a large extent, the universal trends in the VOT durations, with L1 Polish and L2 French being especially consistent in the exhibited patterns. L3 English, on the other hand, demonstrated slightly less consistency in the vocalic effects, thus pointing potentially to a somewhat lower stability of the phonological system in this language. Figures 88-90 present the box plots of the mean VOT values for various vocalic contexts in the three respective languages.

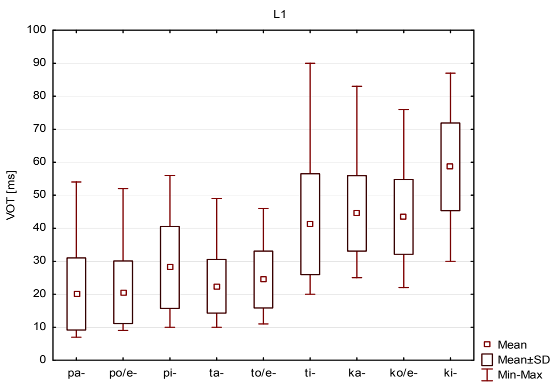


Figure 88. Box plot of mean VOT values for various vocalic contexts in L1 Polish

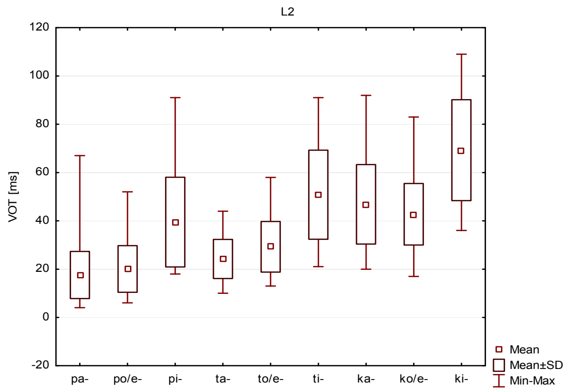


Figure 89. Box plot of mean VOT values for various vocalic contexts in L2 French

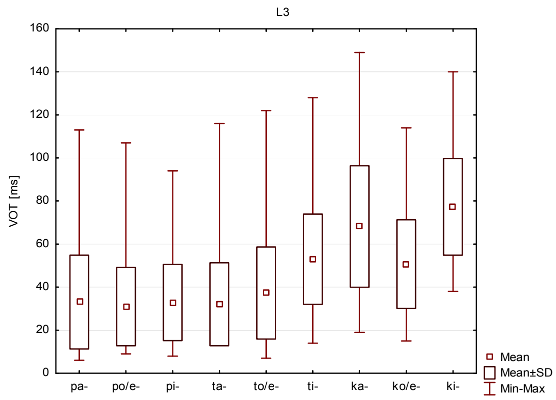


Figure 90. Box plot of mean VOT values for various vocalic contexts in L3 English

5.4.4.10. Analysis of variance

The performed analysis of variance included a two-factor ANOVA between the languages (L1, L2, L3) and the VOT durations of the voiceless plosive sounds /p, t, k/ in Group D.

The results show that the differences in the VOT values within the factors of the language (F (2; 1513) =61.4, p<.05) and the plosives (F (2; 1513) =352.5, p<.05) were found to be significant. Moreover, the interaction between the languages and the plosives on the VOT values was shown to be significant (F (4; 1513) =2.5, p<.05). The findings indicate that there are significant differences between the VOT values for at least two languages, and for at least two plosives. The existing interaction between the languages and the plosives depends on the type of language (L1, L2, L3) as presented in Table 143 and Figure 91.

Table 143. Two-factor analysis of variance for /p, t, k/ and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	61.40 (0.05;2;1513)	0.000000*
plosive (/p/, /t/, /k/)	325.52 (0.05;2;1513)	0.000000*
Language*plosive	2.46 (0.05;4;1513)	0.043507*

*p<0.05

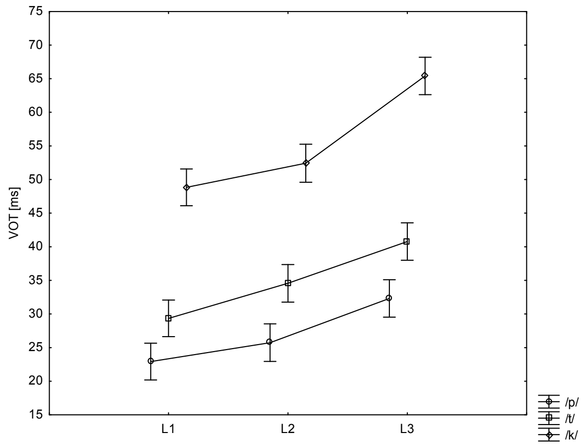


Figure 91. Interaction between the language and plosive factors

Another analysis investigated the interaction of the following vowel context and the language on the observed VOT durations in Group D. To this end, a two-factor analysis ANOVA was performed for the factors of the language (L1, L2, L3) and the context of the vowel following the voiceless plosives in the stressed onset positions in the target words (i.e. low _/a/, high _/i/, and mid _/e, o/). The results demonstrated that there were significant differences in the VOT values within the factor of the languages (F (2; 1513) =48.9, p<.05) and the vowel context (F (2; 1513) =103.3, p<.05). It can be interpreted that significant differences between the VOT values were found for at least two languages, and for at least two vowel contexts. Furthermore, a significant interaction was found between the two factors (F (4; 1513) =5.2, p<.05), which depends on the type of the language (see Table 144, Figure 92).

Table 144. Two-factor analysis of variance for vowel context /-a-/, /-o/e-/, /-i-/ and L1, L2, L3

Factor	F	p
Language (L1. L2. L3)	48.95 (0.05;2;1513)	0.000000*
Vowel (/-a-/. /-o/e-/. /-i-/)	103.33 (0.05;2;1513)	0.000000*
Language*Vowel	5.20 (0.05;4;1513)	0.000369*

*p<0.05

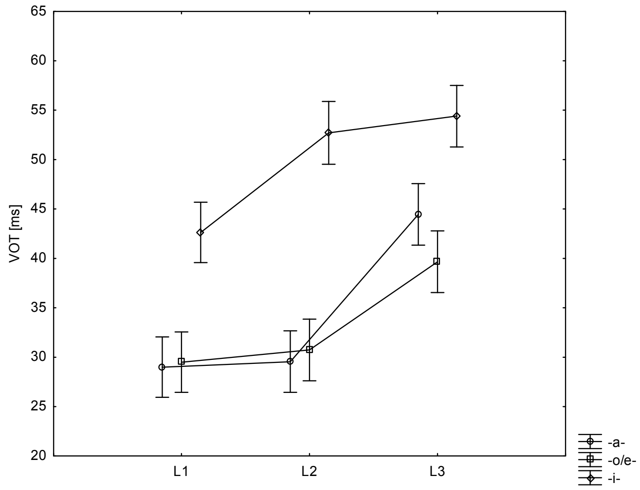


Figure 92. Interaction between the language and vowel context factors

The observed patterns of interaction were, to a large extent, consistent with the universal effects of the vocalic context on the preceding plosive. The high vowel context generated the longest VOT values in all the languages. The differences in durations between the mid and low vowel context were negligible in the case of L1 Polish and L2 French, whereas in L3 English the pattern was slightly different than the expected one, with the mid vowel context /-o/e-/ reported to be shorter than the low vowel context /-a-/.

In order to investigate the language proficiency effects on the generated VOT values, the participants in Group D were divided into two proficiency subgroups, including the 1EG group (lower proficiency) and the 2/3/4EG group (higher proficiency). The results of the conducted two-factor ANOVA for the proficiency groups and for the L1, the L2, and the L3 demonstrated that the differences in the VOT values within the factors of language (F (2; 1516)=32.4, p<.05) were significant. Nevertheless, no significant differences were found between the mean VOT values with respect to the proficiency group (F (1; 1516) =0.7, p>.05) and there was no significant interaction between the factors of the language and the group on the mean VOT values (F (2; 1516) =2.4, p>.05) (see Table 145 and Figure 93).

Table 145. Two-factor analysis of variance for proficiency groups and for L1, L2, L3

Factor	F	p
Language (L1, L2, L3)	32.44 (0.05;2;1516)	0.000000*
Group (1EG, 2/3/4EG)	0.70 (0.05;1;1516)	0.403933
Language*Group	2.42 (0.05;2;1516)	0.089526

*p<0.05

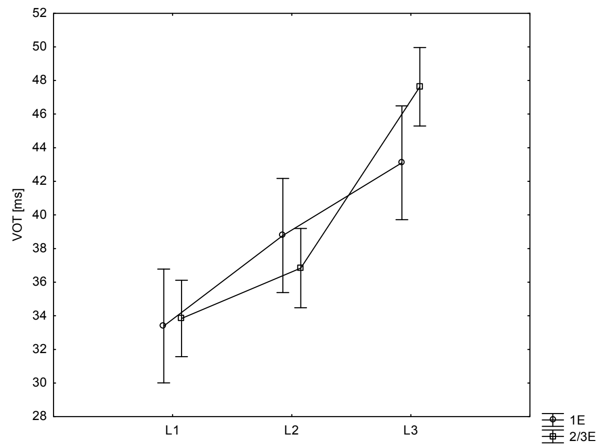


Figure 93. Interaction between the language and proficiency group factors

The results of the two-factor analysis of variance ANOVA for the proficiency groups and the place of articulation PoA /p, t, k/ performed separately for languages L1, L2, L3 show the following patterns. For L1 Polish and L2 French, there are statistically significant differences between the mean VOT values for at least two places of articulation for L1 (F (2; 516) =156, p<.05) and for L2 (F (2; 490) =83.7, p<.05); there are no significant differences in VOT for the proficiency groups; there are no significant interactions between the PoA and the proficiency groups.

For L3 English, there are statistically significant differences between the mean VOT values for at least two places of articulation (F (2; 498) =80.1, p<.05); there are statistically significant differences between the mean VOT values for the proficiency group (F (2;498) =4.3, p<.05); there are no significant interactions between the PoA and the groups (see Table 146 and Figure 94).

Table 146. Two-factor analysis of variance for proficiency groups and place of articulation PoA /p, t, k/ and for languages L1, L2, L3

Factor	L1		L2		L3	
	F	p	F	p	F	p
PoA (/p/. /t/. /k/)	155.96 (0.05;2;516)	0.000000*	83.68 (0.05;2;490)	0.000000*	80.15 (0.05;2;498)	0.000000*
Group (1E. 2/3E)	0.13 (0.05;1;516)	0.718796	1.01 (0.05;1;490)	0.316421	4.34 (0.05;1;498)	0.037640*
PoA* Group	0.10 (0.05;2;516)	0.909157	0.18 (0.05;2;490)	0.832668	0.32 (0.05;2;498)	0.727684

*p<0.05

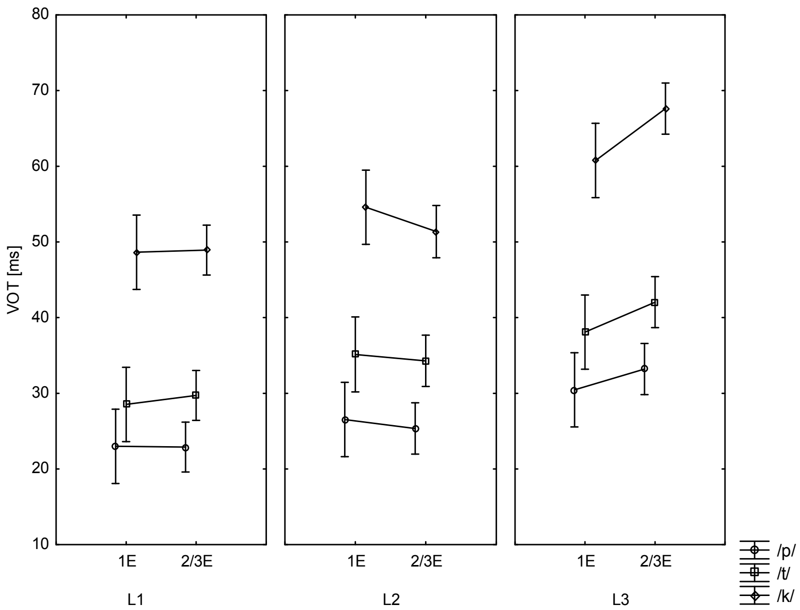


Figure 94. Interaction of PoA (/p/, /t/, /k/) and Group (1E, 2/3E) for languages L1, L2 and L3

5.4.4.11. Multiple regression analysis

Multiple regression analyses were performed to estimate the relationship between the dependent and independent variables in Group D.

The first regression analysis investigated the relationship between the VOT values as the dependent variable and a number of independent variables including the language (L1, L2, L3), the place of articulation PoA (/p/, /t/, /k/), the vowel context and the proficiency group. The assumed levels of significance point to a statistically significant interdependence between the VOT values and the following variables of the language, the place of articulation and the vowel context, whereas the influence of the group variable was not found to be significant. The R squared result indicated that 45.1% of the variance is accounted for by the independent variables (see Table 147).

The beta coefficient was calculated to establish which of the independent variables had a greater effect on the dependent variable in the multiple regression analysis. The beta coefficient values suggest that PoA 1, i.e. /p/ (b=0.44), Vowel 2 (-o/e-) (b=0.21) and Language 1, i.e. L1 (b=0.2) had the biggest influence on the VOT values, whereas the impact of Vowel 1 (-a, -i), PoA 2 /k/, /t/ and Language 2 (L2, L3) was relatively smaller (see Table 148).

Table 147. Multiple regression analysis 1

Variable	R	R^2	F	p
VOT	0.67	0.451	177.75 (0.05;7;1514)	0.000000*

*p<0.05

Table 148. Results of multiple regression 1 for variables

Effect	t	p	Beta
Target word	84.45	0.000000*	–
Language 1	-9.11	0.000000*	-0.200
Language 2	-2.39	0.016744*	-0.053
Group	-1.22	0.222697	-0.023
PoA 1	-20.07	0.000000*	-0.441
PoA 2	-7.01	0.000000*	-0.154
Vowel 1	-8.01	0.000000*	-0.176
Vowel 2	-9.72	0.000000*	-0.213

*p<0.05

A further multiple regression analysis was performed for the L3 VOT values as the dependent variable and the L1 VOT and the L2 VOT as the independent variables. The analysis pointed to a statistically significant interdependence among the variables under investigation. The regression model is explained in 33.9% by the independent variables L1 VOT and L2 VOT (R^2=0.339). The bigger impact on the dependent variable L3 VOT is exerted by the L2 VOT values (b=0.37) than the L1 VOT (b=0.28) (see Tables 149 and 150).

Table 149. Multiple regression analysis 2

Variable	R	R^2	F	p
L3 VOT	0.58	0.339	126.14 (0.05;2;493)	0.000000*

*p<0.05

Table 150. Results of multiple regression 2 for variables

Effect	t	p	Beta
Target word	5.83	0.000000*	–
L1 VOT	6.58	0.000000*	0.287
L2 VOT	8.55	0.000000*	0.373

*p<0.05

As a final step, a regression analysis was calculated for the L3 VOT values as the dependent variable and other independent variables including the L1 VOT, the L2 VOT, the place of articulation /p, t, k/, the vowel context and the proficiency group. The results show that the regression model is explained in 38.7% by the independent variables. The assumed level of testing probability points to a statistically significant impact of such independent variables as the L1 and L2 VOT, the place of articulation /p, t, k/ and Vowel 1, i.e. /-a, -i/. The greatest impact on the L3 VOT is exerted by the L2 VOT (b=0.37) and PoA /t/ (b=0.24), whereas it is smaller in the case of the L1 VOT, the vowel context /-a, -i/. and the group PoA /p, k/ (see Tables 151 and 152).

Table 151. Multiple regression analysis 3

Factor	R	R^2	F	p
L3 VOT	0.62	0.387	44.00 (0.05;7;488)	0.000000*

*p<0.05

Table 152. Results of multiple regression 3 for variables

Effect	t	p	Beta
Target word	5.67	0.000000*	–
L1 VOT	4.35	0.000017*	0.210
L2 VOT	7.66	0.000000*	0.373
Group	-1.03	0.303237	-0.041
PoA 1	-2.91	0.003741*	-0.141
PoA 2	4.90	0.000001*	0.242
Vowel 1	3.45	0.000615*	0.150
Vowel 2	-0.92	0.357615	-0.039

*p<0.05

5.5. VOT goodness of fit – joint analysis

In the final round of analyses, the goodness of fit to the control VOT values was correlated with other L3 performance measures such as (1) the MPhA composite score, (2) accentedness ratings, and (3) the participants' variables.

The performed joint groups correlation analyses between the MPhA composite score and the VOT goodness of fit demonstrated no statistically significant correlations either for the L3 VOT goodness of fit to the L3 controls or for the L2 VOT goodness of fit to the L2 controls (see Table 153).

Table 153. Correlation between MPhA and VOT goodness of fit – joint analysis

Pair of variables	N	Rs	t(n-2)	p
Control L2 & Composite score	105	0.13	1.35	0.179499
Control L3 & Composite score	105	-0.08	-0.78	0.436987

As far as the joint groups correlation analysis between the VOT goodness of fit and the accent ratings is concerned, no statistically significant correlations were found for any of the rating parameters (i.e. foreign accent, comprehensibility or pronunciation correctness) either for the L3 VOT goodness of fit to the L3 controls or for the L2 VOT goodness of fit to the L2 controls (see Table 154).

Table 154. Correlation between VOT goodness of fit and accent ratings

Pair of variables	N	Rs	t(n-2)	p
Control L2 & Accentedness	84	0.13	1.20	0.232056
Control L2 & Comprehensibility	84	0.15	1.40	0.164897
Control L2 & Correctness	84	0.18	1.66	0.101530
Control L3 & Accentedness	84	0.03	0.30	0.763676
Control L3 & Comprehensibility	84	0.12	1.11	0.269566
Control L3 & Correctness	84	0.20	1.85	0.068110

Finally, a joint groups correlation analysis was performed between the VOT goodness of fit and the participants' variables. For the L3 VOT goodness of fit to the L3 controls, several weak to medium correlations were found to hold ($p<0.05$) as it correlated with the years of formal training in the L3 (inverse correlation Rs=-0.42); the age of learning of the L3 (Rs=0.38); the age of learning of the L2 (inverse correlation Rs=-0.41); the years of formal training in the L2 (Rs=0.32); and the participants' age (inverse correlation Rs=-0.22). For the L2 VOT goodness of fit to the L2 controls only one weak correlation with the age of learning of the L3 was observed (Rs=0.18, $p<0.05$) (see Table 155).

Table 155. Correlations between VOT goodness of fit and participants' variables

Pair of variables	N	Rs	t(n-2)	p
Control L2 & L2_YFT	125	0.13	1.45	0.149699
Control L2 & L2_AOL	125	-0.08	-0.91	0.366765
Control L2 & L3_YFT	125	-0.17	-1.90	0.059466
Control L2 & L3_AOL	125	**0.18**	2.04	**0.043124**
Control L2 & N_TOTAL_Ln	125	0.06	0.64	0.525651
Control L2 & Stay in L3	125	-0.06	-0.62	0.538001
Control L2 & Age	125	0.11	1.24	0.217677
Control L2 & evaluat L3 comp	125	0.05	0.50	0.618044
Control L2 & evaluat L3 pron	125	-0.07	-0.80	0.424002
Control L2 & L2_Prof	125	0.13	1.43	0.156558
Control L2 & L3_Prof	125	0.01	0.12	0.907701
Control L2 & Year	125	0.14	1.53	0.129446
Control L3 & L2_YFT	125	**0.32**	3.79	**0.000234**
Control L3 & L2_AOL	125	**-0.41**	-4.98	**0.000002**
Control L3 & L3_YFT	125	**-0.42**	-5.16	**0.000001**
Control L3 & L3_AOL	125	**0.38**	4.50	**0.000015**
Control L3 & N_TOTAL_Ln	125	0.16	1.74	0.084044

Control L3 & Stay in L3	125	0.05	0.50	0.616748
Control L3 & Age	125	**-0.22**	-2.54	**0.012369**
Control L3 & evaluat L3 comp	125	-0.15	-1.74	0.085116
Control L3 & evaluat L3 pron	125	-0.04	-0.46	0.649280
Control L3 & L2_Prof	125	0.06	0.68	0.500407
Control L3 & L3_Prof	125	-0.17	-1.91	0.058045
Control L3 & Year	125	-0.11	-1.17	0.243076

5.6. Discussion

The major aim of this part of the study was to explore the interactions between three phonological systems of the multilingual participants based on their productions of laryngeal contrasts, and, specifically, to investigate the sources and directions of cross linguistic influence in voice onset time patterns. To this end, VOT values in the participants' first, second and third languages in four different groups were measured acoustically and compared to one another as well as to the reference values of the respective languages. The subsequent discussion will address the research questions that were posed in this part of the study.

RQ 1: Do multilingual subjects distinguish between their language systems (i.e., L1, L2 and L3) with respect to the VOT values?

The first research question dealt with whether multilingual participants differentiate between their first, second and third language systems with regard to VOT values. The findings demonstrated that the subjects distinguished between the VOT length in all their respective language repertoires in Group A (L3 French_L2 English) and Group B (L3 German_L2 English). In the other two groups they produced voiceless plosives in stressed onset positions with significantly different values in part of their language systems; in Group C (L3 English_L2 German) no significant differences were found between L2 German and L3 English with respect to /t/ and /k/, yet all the remaining differences were significant; and in Group (D L3 English_L2 French) there were no significant differences between L1 Polish and L2 French for /p/ and /k/, with all the remaining VOT differences being significant. Further, the analyses of variance for all the groups indicated significant interactions between the factors of language (L1, L2, L3) and the VOT durations in voiceless plosives /p, t, k/.

In the majority of cases, foreign language categories proved sufficiently dissimilar acoustically from the established inventory of L1 Polish phonetic categories for the subjects to modify their realizations of /p, t, k/ in their foreign languages (i.e. L2 and L3) so that the implementations of voiceless plosives differed significantly across all the language systems. The only exception to the dissimilarity between L1 VOT values and VOT values in other foreign languages was Group D, where no significant differences were found between L1 Polish and L2 French for /p/ and /k/; however, this could have been expected taking into account similarities between VOT durations in Polish and French as voicing languages.

Interestingly, in three out of four groups the L3 VOT intervals were not assimilated either to the L1 or the L2 values except for Group C, where the L3 English values approximated those in L2 German. It appears that in Groups A, B, and D a modified category, different from the L1 and L2 VOT systems, was formed for /p, t, k/ in the respective third language by adding different realizations thereof. It, thus, follows that the participants showed some evidence of restructuring of their phonetic space. This ability of learning new patterns of segmental articulation did not seem to diminish after a critical period as all the participants were late learners. This finding is in line with Flege's (1995) Speech Learning Model (SLM), which posits, among others, that the phonetic system of a learner remains adaptive throughout their lifetime and open to modifications of the phonetic categories.

As far as the VOT performance in the first and second language systems is concerned, some interesting patterns were observed in the studies. The mean values for the L1 Polish stops /p, t, k/ did not differ across the studies and fell within the established literature reference VOT values for the Polish monolinguals. However, some VOT lengthening was observed compared to the referenced Polish norms, particularly in Group B in which VOT durations in L1 Polish were significantly higher than the reference values, yet still within the acceptable 5 ms range. This slight lengthening of native Polish VOT durations could be attributed to the influence of the established long-lag values in L2 English (Group B). This phenomenon can be interpreted as an evidence for bi-directional cross-linguistic influence or, the so called, 'regressive transfer', as attested earlier in the SLA literature (e.g., Flege, 1987; Waniek-Klimczak, 2011). On the whole, the observed VOT patterns in L1 Polish approximated very closely the reference literature values (Keating et al. 1981) thus lending external validity to the conducted study and the obtained VOT measurements on the one hand, and, on the other, confirming Keating et al.'s (1981) observations concerning VOT values in Polish. Moreover, the findings for L1 Polish indicate that native VOT values in multilinguals remain more or less stable in spite of the acquisition of further language systems with varying VOT durations, and that they maintain a resemblance to the monolingual Polish norms as long as the first language remains the dominant language, and the second and third languages retain their respective status. It should be further emphasised that the participants were not early bilinguals or multilinguals, but the acquisition of foreign languages occurred at a later stage in life through formal classroom instruction rather than naturalistic exposure, which could have also had an impact on the relative stability of their first language system with respect to the VOT.

The assumption of a facilitative L1 transfer on the acquisition of French VOT patterns in Groups A and D was not substantiated in the analysis of the data, as could have been expected. By means of transferring the L1 Polish VOT values into their L3 or L2 French, the participants in groups A and D should have been able to produce French voiceless stops in a target-like manner, yet this was not found to be the case. The observed intermediate values for the French VOT duration suggest the existence of some additional intervening variables, other than the L1 influence on the patterns of VOT acquisition.

As far as the VOT patterns in the second language are concerned, in three out of four groups, the participants approximated quite closely their L2 VOT values to the respective target durations and native controls (Groups A, B, D). For L2 English in Groups A and B, the observed VOT durations were somewhat higher than the literature reference norms, but still within the acceptable 5-10 ms range, and very close to the control English VOT measurements (with on average a 2 – 6% goodness of fit ratio). In the case of Group C with L2 German, some conflicting results were reported, as the L2 VOT values approximated the selected literature reference values for German (cf. Angelowa and Pompino-Marschall, 1985), yet they fell below the control German measurements (with on average a -30% goodness of fit ratio).

On the whole, it may be concluded that the VOT values in the second language approximated quite successfully the respective target norms, much more closely than in the case of the L3 VOT durations. This observation can be attributed to the participants' high proficiency in the respective second languages as well as to several other factors, such as an early age of onset for L2 acquisition, a long exposure to the target language, the recency and intensity of L2 use and phonetic training in this language. Particularly noteworthy is the extent to which the L2 VOT values in Groups A and B approximated the target English native norms in comparison to the related studies on the L2 English acquisition of VOT patterns by Polish advanced learners, as reported in the SLA literature (e.g. Waniek-Klimczak, 2011).

RQ 2: Do the L3 VOT patterns approximate the participants' values in their L1, or L2 or the L3 native norms?

The second research question focused on the VOT values in the third language and was aimed to explore whether the L3 VOT patterns approximate the participants' L1, or L2 or the L3 native norms. The reported L3 VOT values in all the groups (A, B, C, D) were off the target; approximately 30% below the native English controls in the case of L3 English in groups C and D; approximately 20% below the German controls for L3 German (group B); and approximately 25% above the French controls for L3 French in group A.

In general, the observed L3 VOT durations constituted 'hybrid' values, following two established patterns:

1) Compromise L3 VOT values between the L1 and L2 (as in groups A and B)
– in group A – the L3 French VOT was intermediate between the L1 and L2, i.e., longer than the L1 Polish values (and the French controls), but shorter than those for L2 English,
– in group B – the L3 German VOT was intermediate between the L1 and L2, i.e., longer than the L1 Polish values, but shorter than those for L2 English (and the German control VOT durations).

In the case of these groups, the L1 and L2 seemed sufficient to account for the observed VOT patterns in the L3.

2) Compromise L3 VOT values between L1/L2 and the target norms (as in groups C and D)
– in group C – the L3 English VOT was intermediate between L1 VOT and L2 VOT / English norms, i.e. longer than in L1 Polish, similar to L2 German values, but shorter than the typical target English norms
– in group D – the L3 English VOT durations were intermediate between L1/L2 VOT and the English target norms, i.e. longer than the L1 Polish and L2 French values, but shorter than the typical English VOT values.

In these groups, it was not possible to tease apart the potential impact of the L3 target values as the control VOT durations were too close to the L1 Polish values (in the case of group A with L3 French) or too close to L2 English (in the case of group B with L3 German). The L1 and L2 systems alone were thus not sufficient to account for the obtained L3 VOT measurements, and some influence of the target values of the respective third language was attested.

If the participants had identified the foreign language sounds in terms of the native L1 categories, they would have used the established articulatory patterns for the production of the L3 voiceless plosives, i.e. /p, t, k/ would have been implemented identically as short lag in L3 French, English and German, yet the obtained data do not confirm this. The participants made articulatory modifications in their realizations of initial voiceless plosives in their L3s, either under a potential influence of the long-lag VOT categories established for L2 English (as in groups A and B), or in trying to approximate the respective norms of the third language (as could have been the case in groups C and D).

In accordance with Flege's (1995) Speech Learning Model SLM, it was hypothesized that the VOT patterns in L3 French (group A) would approximate the L1 Polish phonetic norms because of the similar realization of the voiceless/voiced plosives distinction, due to the phenomenon of equivalence classification. However, the results failed to confirm the expected facilitative transfer of the L1 Polish VOT values into L3 French. On the contrary, it was demonstrated that the target words in L3 French were produced with mean VOT values that were intermediate to the phonetic norm for VOT in L1 Polish and L2 English. The most probable interpretation of the results is that the established L2 English VOT values exerted some influence on the acquisition of the additional foreign language, as the values observed for the L3 were longer than the typical native Polish and native French VOT intervals. A similar pattern was observed in group B, thus a parallel interpretation can be applied to the L3 German VOT, with the additional potential impact of the German target norms reinforcing the L2 English effect. The L2 was also found to influence the production of /p, t, k/ in L3 German (Group B), although in this case no L1 positive transfer was hypothesized, as the VOT values typical for native Polish and German speakers differ significantly.

In the case of groups C and D (with L3 English), the L1 and L2 VOT values alone were not sufficient to account for the L3 VOT patterns, as the respective L1 and L2 VOT intervals were characterized by a shorter lag than the target English norms. The L3 values in Group D did not undergo equivalence classification under the impact of L1 Polish and L2 French VOT durations, but demonstrated lengthening properties – which may point to a potential awareness of the longer lag target English VOT. Similarly, in Group C, the L3 English VOT values were closer to those in L2 German than in L1 Polish. The findings pointed to compromise or 'hybrid' VOT values in the L3, which were intermediate between the existing systems in the first and second language and the target norms for English, which was the third language.

It appears, referring to Flege's terminology, that the L3 voiceless plosives were most probably categorized as 'similar' but not 'identical' to the Polish /p, t, k/, as they were implemented with a different lag length in all the respective third languages in groups A, B, C and D than their counterparts in the native Polish sound system. A lower proficiency level in the L3s and a relatively short period of exposure to the L3 prevented the learners from establishing more target-like native VOT values in the third language, although they achieved this goal in their more advanced L2s.

RQ 3: Is there a proficiency group effect on L3 VOT measurements based on the amount of L3 exposure?

The third research question investigated if there was any group effect on the L3 VOT measurements based on the length of exposure to the third language. The analysis of the proficiency group effects was based on statistical testing and provided mixed results. There were two apparent trends observed in the data:

1) no significant L3 proficiency group effects (as in Groups B and D)
2) significant or partially significant L3 proficiency group effects (as in Groups A and C).

In the first case, in Groups B and D neither the conducted Mann-Whitney test for the mean VOT in /p, t, k/ in the three language systems for two proficiency groups, nor the performed two-factor ANOVA for the proficiency groups and L1, L2, and L3 pointed to any significant differences in the VOT values with respect to the L3 proficiency group. Further, there was no significant interaction between the factors of the language and the group. Finally, the multiple regression analyses performed to investigate any relationships between VOT as the dependent variable and the number of independent variables did not show any interdependence between the proficiency group variable and the observed VOT patterns. Consequently, in Groups B and D no regularities were observed in this respect in the VOT durations in all the languages involved, thus the L3 proficiency level cannot be regarded as a significant predictor of the VOT intervals. However, the lack of a significant difference between the performance of the various proficiency subgroups in Groups B and D should not be regarded as a complete disconfirmation of the hypothesis that the proficiency group is a predictor of the VOT performance in the L3 and that the multiliguals' L3 interphonology may be subject to further modifications. The scope of the L3 proficiency levels may have not been sufficient enough to detect significant group effects, therefore, more diverse groups might be necessary to ascertain a difference.

On the other hand, in Groups A and C, significant or partially significant effects of the L3 proficiency group were reported. The effect of the group proved strongly significant for Group A in all of the conducted statistical tests (i.e. the Mann-Whitney test, ANOVA and the regression analysis) and partially significant for Group C in some of the performed tests. In Group A the observed trends corresponded to the expected proficiency group effects as the more advanced group exhibited VOT values that were closer to the target durations in L2 English and L3 French. Conversely, in Group C, the Mann-Whitney test demonstrated some differences in VOT durations as conditioned by the L3 proficiency, but there were no consistent regularities apart from a tendency in L2 German for a longer lag VOT in the more advanced group. This was partially confirmed by an analysis of the variance for the proficiency group and the place of articulation (PoA); however, the multiple regression analysis showed no interdependence between the group variable and the VOT values.

Recapitulating, the difference in the L3 proficiency level was not found to be a strong predictor of the VOT production across the different language pairing groups. However, an adaptation towards target-like VOT values that was observed to occur in parallel with the degree of advancement in L3 in Groups A and C provides some limited support for the proficiency group effect. Although there is mixed evidence that the length of exposure to the L3 significantly influences the degree of approximation to the native target norm, there were some indications pointing to this effect.

RQ 4: Which other factors have an impact on VOT values in the three languages?

The fourth research question explored factors that influenced the VOT values in the multilingual participants' language repertoires.

The cross-linguistic correlations between the VOT values for /p, t, k/ in the L3 vs. L1, and the L3 vs. L2 language pairings, pointed to significant positive correlations in all the groups (A, B, C, D); however, they differed in magnitude. In the majority of groups (B, C, D), the cross-linguistic correlations demonstrated a stronger covariance between the L3 and L2 VOT patterns; with moderate to strong correlations between L3 German and L2 English in Group B, somewhat weaker correlations between L3 English and L2 German in Group C; and the weakest covariance between L3 English and L2 French in Group D. Anyway, in all these analyses, L1 Polish correlated to a lesser degree with the respective L3. The results were confirmed by the multiple regression analyses for the L3 VOT values as a dependent variable, which pointed to a considerably stronger effect of the second rather than the first language VOT as independent variables in groups B, C, and D. In Group A, the covariance between the variables remained weak to moderate, yet Pearson's correlations between the VOT patterns in L3 French and L1 Polish tended to be higher for all the voiceless plosives than between L3 French and L2 English. This trend was confirmed by the results of the multiple regression analysis for the L3 VOT values as a dependent variable, which demonstrated a stronger impact exerted by the first language rather than the second language VOT as independent variables. The general regression analyses for the VOT values in Groups A, B, C, and D indicated that among the independent variables, including the language (L1, L2, L3), the place of articulation, the vowel context and the proficiency group, the ones that proved to exert the greatest effect on the dependent variable were the language systems and the place of articulation.

The findings provided evidence for the co-existence of the prevailing L2 effect and some underlying L1 interference as factors that exert an impact on the acquisition of the VOT patterns in the third language. Furthermore, they substantiated the assumption of a combined cross linguistic influence in third language acquisition, as suggested by De Angelis (2007). The present results are mostly consistent with previous studies on L3 phonological acquisition (Llama et al., 2010; Wrembel, 2010, Wunder, 2010), which pointed to a combined CLI from both the native and non-native languages. The results contradict the findings by Ringbom (1987) or Pyun (2005), who observed a prevailing influence of the L1 phonology on L3 acquisition.

RQ 5: Does the typological proximity between particular language repertoires influence the reported VOT patterns?

The next research question was concerned with whether a typological proximity between language repertoires influenced the VOT values in L3 acquisition. To this end, patterns of L3 VOT acquisition were compared across groups with greater typological proximity with respect to the VOT parameter, i.e. Groups B and C (L3 German_L2 English and L3 English_L2 German) and greater typological distance, i.e. Groups A and D (L3 French_L2 English, L3 English_L2 French). In order to verify the typology effect, one would expect a parallel behaviour in groups with the same typological profile.

Interestingly, although similar patterns of the VOT performance were found in the typologically related groups (B and C), it was not the case for the typologically less related groups (A and D). In Group B (L3 German_L2 English), the regression analysis demonstrated strong L2 effects on the L3 VOT with a limited effect of the L1. A parallel tendency was found in Group C with a reverse language pairing (L3 English_L2 German). In these two groups it is difficult to tease apart the L2 status effect from the potential impact of typology, and it is impossible to state univocally which one is the actual predictor of the L3 VOT performance. This could be interpreted as a combined effect of the L2 reinforced by the typological closeness between these two languages, especially with respect to the VOT parameter, since both English and German are classified as aspirating languages with a long lag VOT.

Nonetheless, the typological effect was not confirmed in the two remaining mirror groups (A and D) with less typologically related language pairings (L3 French_L2 English and L3 English_L2 French), which did not exhibit a parallel behaviour. Although the regression analysis in Group A demonstrated, as expected, little impact of L2 English on the L3 French VOT, there was a stronger indication of such an influence of L2 French on L3 English in Group D. On the other hand, a considerable impact of L1 Polish on the L3 French VOT was evidenced in Group A, thus pointing to a typological effect, as Polish and French are both voicing languages characterised by a short lag VOT, which makes them typologically related languages with respect to the VOT parameter.

Further comparisons were conducted across the groups with a similar typological distance. It resulted in yet another counterargument against the effect of typology since there were striking similarities between the mean VOT durations for /p, t, k/ in L3 French (Group A) and L3 German (Group B), irrespective of the typological proximity between the language combinations involved in the groups. If typology was at stake, Group B should be more influenced by L2 English than Group A, but this was not the case, and thus the L2 status seems to be a stronger predictor than typology. On the other hand, in groups with L3 English and different L2s (Groups C and D), the mean VOT durations for voiceless plosives were nearly identical as well, irrespective of the fact whether the L2 was German – longer lagged and more typologically related (Group C) – or French – with shorter VOT values and a less typologically related status (Group D). The reported similarity of the L3 English mean VOT measurements in Groups C and D does not confirm either typology or the L2 status as a determining factor for the L3 acquisition patterns of VOT.

Recapitulating, no conclusive evidence of the typology effect can be drawn from the analysis. A partial explanation for the approximation to the native VOT norms in Groups B and C may be attributed to the close typological proximity between English and German, yet one cannot tease it apart from the L2 effect. Moreover, the results of the conducted regression analysis pointed to a higher percentage of the variance that was accounted for by the influence of the L2 variable on the L3 VOT rather than the first language variable in three out of the four groups. A partial explanation of these results may be a closer typological proximity between the pairs of languages involved with respect to the VOT dimension. The impact of L1 Polish seemed more noticeable in the case of L3 French (Group A) as these two languages are the so called 'voicing' languages which make a distinction between voiced and voiceless unaspirated stops, whereas the effect of the second language on the L3 prevailed in English-German language pairings (Groups B and C) since these two languages can be categorized as 'aspirating' languages which distinguish between voiceless aspirated and voiceless unaspirated stops (cf. Lisker and Abramson, 1964).

RQ 6: Do VOT patterns in trilingual acquisition resemble those reported in the SLA literature or are these trends specific for each context of acquisition?

The final research question aimed to investigate whether the VOT patterns in third language acquisition resemble the ones reported in SLA studies. On the basis of the conducted analyses, the category assimilation observed in the L3 VOT values varies from the one reported in the SLA literature, where it is commonly attested to occur between the L1 and L2 categories, thus forming a hybrid between the native and target values (e.g., Flege, 1987; Flege and Eefting, 1988). In third language acquisition this compromise is of a different, more complex nature because of the co-existence of three language systems in the multilingual participants' minds.

In all the groups (A, B, C, D) the L3 VOT values were found to deviate both from the native Polish and the target native norms for the respective L3s. There was some evidence of an intervening variable of the L2 as the observed L3 values were usually found to be intermediate between those of the native language (L1 Polish) and the previously acquired foreign language (L2), except for Group D. The potential impact of the target values of the additional foreign language (L3) seemed to have been also attested, especially in Groups C and D.

On the whole, the observed patterns of VOT acquisition in the third language resembled those reported in the SLA literature in the sense that they also involved a hybrid value. However, their specificity was related to the fact that these hybrid VOTs were not limited to a compromise between the native and the target VOT intervals, but accounted for an intervening variable of the previous foreign language system known to the multilingual participants.

Finally, the analysis looked at whether the VOT patterns in third language acquisition follow the universal effects of the place of articulation (PoA) and the vocalic context, as reported widely in the SLA studies. The findings confirmed a general adherence to the universal principles in nearly all the language systems in all the groups under investigation (A, B, C, D). The results demonstrated progressively longer VOT values for velars when compared to alveolars and bilabials, a trend universally attested by Maddieson (1997). As far as the adjacent vocalic context is concerned, the VOT durations tend to be longer when a plosive is followed by a high rather than a low vowel. On the whole, the vocalic context effects observed in all the groups follow, to a large extent, the universal trends in the VOT durations in the preceding plosives, but this trend was especially consistent in the first and second languages, whereas the respective third languages exhibited less stability in this respect (Groups B, C, D). This may come as no surprise given the well established L2 systems and significantly less proficient L3 systems in all the groups. By means of demonstrating the expected language universal effects with regard to the VOT related phenomena, the external validity of the studies was confirmed.

5.7. Conclusion

In conclusion, this part of the series of studies aimed to shed more light on the growing area of L3 phonological acquisition by identifying unique interlanguage VOT patterns and investigating the complex interaction between several phonological systems in the multilingual learners. The contribution was expected to provide new insights into the phenomenon of cross linguistic influence in the trilingual acquisition of the voice onset time patterns. The major finding was that the multilingual participants differentiated between their respective language systems, i.e., L1, L2 and L3 with respect to the VOT durations. Moreover, they usually created new 'merged' VOT categories for the third language, which deflected away from both the L1 and L2 categories, thus maintaining a phonetic contrast between the three language systems at their disposal.

This series of studies investigating various language combinations and different proficiency groups was also intended to tease apart the effects of language typology and language proficiency on the acquisition patterns in third language phonology. To this end, it was designed as a large cross-sectional study involving comparable parallel groups in order to safeguard that valid generalizations may be drawn from it. On the whole, a combined cross-linguistic influence was observed as the phonetic properties under examination were generally transferred from L1 Polish, and the respective second languages onto the third language interphonology, thus resulting in the compromise VOT values for the L3. Consequently, the present studies have provided further evidence for the L2 effect in the phonological acquisition of a third language. Conversely, the results have undermined the view that the mother tongue was the only source of potential cross-linguistic influence in multilinguals acquiring another foreign language. Further, the typology was reported to have a facilitating effect on the acquisition of the L3 VOT patterns; however, it was not confirmed to be a decisive predictor of success.

Admittedly, the study suffered from some limitations as the VOT measurements were performed on speech samples in the reading style only. Furthermore, the validity of the monolingual reference values for particular languages may be questioned on the grounds of methodological heterogeneity. However, the observed VOT patterns for L1 Polish (in all the groups) and L2 English (in groups B and C) demonstrated a very close resemblance to the reference values, as reported in the literature. Finally, the L3 proficiency group effect might be more visible had the studies allowed for a more diverse level of samples. Future planned studies will involve a longitudinal design in order to trace more developmental changes and progress over time.

Chapter 6

Study III – Metaphonological awareness in L3

6.1. Introduction to MPhA

The acquisition of phonology in a foreign language has been claimed to involve an array of factors related to the linguistic, psychomotor, and cognitive domains. In spite of the acknowledgements that cognitive processes of awareness and attention strongly influence various aspects of a foreign language including phonology, the cognitive domain has remained largely unexplored as far as the phonological development is concerned (Ellis, 1999; Fraser, 2010; Sicola, 2010). To date, little attention has been paid to how learners attend to their phonological output and modify it. Therefore, the present study aims to bridge the existing gap and shed more light on metalinguistic awareness in the area of foreign language phonology, hereafter referred to as metaphonological awareness (MPhA). Further, the notion will be explored from a multilingual perspective which entails an interaction of metalinguistic awareness with an additional component of cross-linguistic awareness, as stipulated by Jessner (2006). Previous exploratory investigations into metaphonological awareness by the present author were presented in Wrembel (2013, 2015).

6.1.1. Role of awareness in SLA/TLA

Metalinguistic awareness has been recognised by many scholars as an important component of language proficiency which facilitates foreign language acquisition (Alderson et al. 1997, Gombert 1992, Herdina and Jessner 2002). Jessner (2006: 42) defines metalinguistic awareness as “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”. More specifically, James (1999) distinguished between language awareness as such (i.e., language awareness as metacognition reflected in one’s intuitions about language) and consciousness raising (i.e., language awareness as cognition generated through objective linguistic knowledge).

The role of awareness in language learning has been discussed extensively in the second language acquisition literature. The majority of studies have referred to the influential noticing hypothesis put forward by Schmidt (1990, 1993, 2001). In his hypothesis Schmidt maintains that noticing is a precondition for input to be converted into intake and no learning can take place without noticing (1990: 129). Moreover, he stresses the vital role of attention in foreign language learning by claiming that "attention appears necessary for understanding nearly every aspect of second language and foreign language learning" (Schmidt 2001: 6). Schmidt distinguished different levels of awareness, with awareness at the level of noticing and understanding being the most relevant for foreign language acquisition.

Another model advocating awareness or consciousness as a precondition for foreign language learning was proposed by Bialystok (1982, 1994, 2001) and Bialystok and Ryan (1985). Bialystok's Model assumes that language learning involves a process of developing analysed knowledge (i.e. conscious, not intuitive knowledge) and control over that knowledge. In the process of language development, knowledge becomes more analysed and complex. The model stipulates that implicit knowledge becomes explicit and can be subject to conscious analysis. On the other hand, the control component selects and co-ordinates items of information and knowledge as well as enables this performance to become automatic. Furthermore, control involves the access procedures to the representation of knowledge and is responsible for the development of selective attention. The two components of analysed knowledge and control are necessary for different types of processing; i.e. the former is responsible for accuracy, whereas the latter is related to fluency. Following this line of thinking, metalinguistic awareness would require a high degree of both analysed knowledge and processing control.

When analysing different models of second language acquisition of speech (cf. Gut 2009 for an overview thereof), the one that comes to the fore as embracing the role of consciousness is Dziubalska-Kołaczyk's (1990) natural model of acquisition of second language phonology. It stems from the functional framework theory of Natural Phonology, as founded by Stampe (1979). Dziubalska-Kołaczyk demonstrated that second language acquisition of phonology involves conscious and controlled learning, which differs substantially from the mechanisms of automatic and unconscious acquisition of speech operating in the first language. The learner observes and imitates the foreign language output, however this process is conditioned by socio-psychological factors and facilitated by formal instruction. Formal setting learners are expected to apply more foreign language processes consistently in their L2 productions than it is the case for naturalistic acquirers. Ultimately, the process of learning foreign language speech may lead to "a total unsupression and correct limitation of those natural phonological processes of a pre-linguistic stage which were selected to operate in the language learned" (Dziubalska-Kołaczyk 1990: 65). Consciousness constitutes a pivotal role in the proposed account since the 'conscious competence of performance', a term coined by Dziubalska-Kołaczyk (2002), allows learners to suppress the negative impact of their current language system due to the meta-knowledge of what is at their disposal.

One of the further issues that has led to scholarly debates in the field, has been the distinction between intuitive language awareness and metalinguistic knowledge. To address this discrepancy, Gombert (1992) distinguished between epilinguistic and metalinguistic awareness. The former type of awareness tends to be unconscious, spontaneous and contextualised, and it involves e.g. self-repair in speech performance. The latter is characterised as decontextualised, conscious and intentional and it precludes conscious reflection on language properties and metalinguistic analysis.

An overview of studies conducted on the role of awareness in phonological acquisition is presented in section 3.2.9 in Chapter 3. For a more detailed discussion of the role of awareness in second and third language acquisition, see Wrembel (2013, 2015).

Since the constructs of awareness and attention are rather complex, their operationalization and measurement is subject to debates. More traditional measures involve offline methods such as questionnaires, learners' diaries, recall charts or retrospective reports; however, recently more attention has been focused on online methods that feature processing time recording or think-aloud protocols (TAPs), i.e. audible verbalizations of thoughts when completing a task (see Medina, 2008: 14-29 for a detailed overview). The online measures of attention appear superior over the offline methods due to the fact that they are much less subject to memory constraints and potential reconstructive processes. According to Schmidt (2001: 20), verbal protocols provide the best evidence that something has been consciously perceived or noticed. Despite initial criticism that thinking aloud while performing an L2 task may alter the learners' primary cognitive processes, this data collection procedure has been acknowledged as reliable and robust (Leow, 2001) and has been widely employed in SLA investigations in recent years (Medina, 2008). Therefore, an adaptation of TAP has been applied in the present investigation with the view to the collection of data on metaphonological awareness (see section 6.2.2. for a detailed description of the procedure).

6.1.2. Multilingual perspective

The present investigation is aimed to explore metalinguistic awareness from a multilingual perspective enriched by an additional component of cross-linguistic awareness, as suggested by Jessner (2006).

Referring to Bialystok's (2001) model of attention and control, scholars investigating TLA argue that multilingual learners rely on an increased monitoring mechanism which performs the usual control functions in speech production (i.e. using anticipatory, corrective and compensatory strategies), but it also allows the learners to monitor the resources in their linguistic repertoires, e.g. by keeping the systems apart (Herdina and Jessner, 2002). A number of studies have been conducted to test Bialystok's (1994, 2001) claims concerning the superiority of bilinguals with respect to metalinguistic abilities, and some attempts have been made to apply them further to multilingual learners (e.g. Gibson and Hufeisen, 2011). These enhanced metalinguistic abilities involve, on the one hand, control of attention (i.e. detecting violations in the linguistic output by focusing attention on a selected linguistic feature) and, on the other, the analysis of structure (i.e. analyzing linguistic structures and attending to them selectively).

Literature on third language acquisition has identified the relevance of metalinguistic awareness as a major component of multilingual competence and as a major factor facilitating the acquisition of additional languages (cf. Alderson et al. 1997, Bono, 2011; Cenoz, 2003; Herdina and Jessner, 2002; Jessner, 1999; Jessner, 2006; Renou 2001; Ringbom, 1987; Roehr and Gánem-Gutiérrez, 2009). Moreover, it is widely acknowledged that multilingual learners enjoy a strategic advantage for subsequent language learning, which is primarily attributed to enhanced metalinguistic awareness. In an attempt to conceptualise these enhanced language learning skills in multilinguals, scholars assume the existence of a multilingual asset, which is referred to as e.g. 'M-factor', or 'multilingualism factor' (Herdina and Jessner, 2002; Jessner, 2006) or the cumulative positive effect of enhanced MLA (metalinguistic awareness) stemming from Hufeisen's factor model (e.g. Gibson and Hufeisen, 2011).

6.2. Research design

6.2.1. Aims and research questions

This part of the series of studies aimed to explore cognitive processes associated with metalinguistic awareness of phonological performance in a third language. It focused primarily on the aspects of attention, noticing, metacognition and phonological reflection, which are rarely investigated phenomena of foreign language phonological acquisition. Additionally, metaphonological awareness was explored from a multilingual perspective, in order to substantiate a claim that metalinguistic awareness constitutes a significant part of multilingual competence (cf. Jessner 2006).

The investigation was intended as a large-scale follow-up study conducted on four groups with varied language repertoires (Groups A, B, C, D) to further verify the preliminary observations from the author's previous research (Wrembel 2013, 2014) and to embed the findings within a wider context of related investigations.

The study aimed to investigate qualitative and quantitative aspects of metaphonological awareness manifested through the participants' self-repair and modifications of pronunciation mistakes in L3, conscious analysis of their oral performance, self-awareness of problems in L3 pronunciation as well as the level of metacognitive control, and comments on the process of learning L3 pronunciation. Furthermore, one of the major objectives was to explore whether multilingual participants can reflect on their phonological performance in L3 and whether they are aware of whether it is their L1 or L2 that constitutes a source language for cross-linguistic influence in their L3 speech performance.

Specifically, the study was designed to investigate the following research questions:

RQ 1: What forms of metaphonological awareness are manifested by L3 learners? To what extent are the participants able to modify their phonetic output in L3?

RQ 2: How complex is the conducted metaphonological analysis of L3 performance?
RQ 3: Is there evidence of cross-linguistic awareness in L3 learners? What is the source language for CLI in L3 according to the participants?
RQ 4: Is there a correlation between the MPhA composite score and accent rating parameters in L3?
RQ 5: Is there an interdependence between the MPhA composite score and the participants' proficiency level as well as other variables?

6.2.2. Participants and procedures

This empirical part (i.e. Study III) involved the same four groups A, B, C, D as the parallel investigations into foreign accentedness (Study I) and voice onset time (Study II). The number of the participants who took part in the metaphonological awareness protocols was slightly lower than the original group count since some of them did not complete this task due to various reasons. Nonetheless, the number of participants in Study III was fairly balanced across the groups: Group A (N=27), Group B (N=25), Group C (N=29), Group D (N=26). A detailed description of the groups and participants' profiles is presented in section 4.1.3.

The study was designed to be process-oriented in order to supplement quantitative data with a qualitative analysis of the metaphonological awareness, since such complementary analyses of metalinguistic awareness are relatively rare (cf. Roehr, 2006). The research design of the present study was adapted from previous exploratory investigations (Wrembel 2013, 2015) with some further modifications. Metaphonological awareness was investigated through the application of specific verbal protocols, the so called Think Aloud Protocols (TAPs). TAPs constitute introspective research instruments that involve verbalised reports aimed at disclosing the participants' intuitions and mental processes when performing a given task. This method is frequently applied in research on human information processing (Ericsson and Simon, 1984) but also on meta-awareness in multilingualism (e.g. Jessner, 2006) or pronunciation monitoring strategies (Osborne, 2003). Cohen (1006) claims that various aspects of human behaviour can be reflected by means of verbal reports such as:

- self-report, i.e. learners' descriptions and general statements about the learning process,
- self-observation, i.e. inspections of specific language behaviour, both introspective and retrospective,
- self-revelation, i.e. disclosure of thought processes, thinking aloud.

The advantage of verbal reports over other methods aimed at gaining insight into metalinguistic awareness such as self-report interviews and questionnaires is that TAPs elicit introspective data close to the moment of occurrence, and thus do not suffer from memory limitations (Jessner, 2006). With this in mind, think-aloud protocols were adapted as a method of data collection for the purpose of this series of studies; however, they were slightly modified to enable the performance of two parallel tasks, i.e. analytic listening to a recording and the simultaneous verbalisation of comments. In the study the verbalisation was not fully concurrent with the analytical listening but followed with a slight delay of approximately 2-4 seconds. This modification of the TAP procedure was necessary to accommodate the research tasks requirements since simultaneous performance of the two tasks was virtually impossible, as attested in a pilot study. Consequently, the resulting verbal protocols were not fully concurrent, but rather could be classified as stimulated or immediate recall protocols (cf. Gass and Mackey, 2000).

As in the previous exploratory studies (Wrembel 2013, 2015), the data collection procedure involved immediate retrospective and introspective protocols, in which the participants were asked to attend to, modify and comment on their L3 phonological output after listening to short excerpts of their previous text reading recording in this language. The research design consisted of three stages (see Table 156), which were administered as individual sessions in a quiet room. The participants were seated in front of a computer screen with an audio recording program, wearing a headset with a microphone and headphones.

Firstly, the participants were asked to read a short text in their respective L3 English, French or German. The selected texts were about 100 words long and were adapted to the participants proficiency level. The text reading task was recorded by the researcher in a soundproof studio using Audition CS 5.5 as 16-bit mono files at 41000 Hz sampling frequency (Recording 1). The participants had a few minutes to familiarize themselves with the text.

At the second stage, the participants were asked to listen attentively to Recording 1, which was played back through the headphones in short fragments (2-3 seconds long) and to focus on their own pronunciation performance in L3. During the immediate retrospective protocols, which were parallel to the analytic listening, the participants were requested to correct and modify any pronunciation mistakes that they noticed and to comment on their L3 phonetic performance, in an attempt to verbalise immediately any thoughts that came to their minds. When performing the verbal protocols, the participants could resort to any language they liked; however, in the majority they used L1 Polish with minor interjections from their respective second or third languages.

Finally, the follow-up part of the Think Aloud Protocols involved introspective protocols, which consisted in self-reflection on the process of the acquisition of third language phonology and took the form of a semi-structured interview. The interaction with the researcher was limited to an occasional provision of support questions. The retrospective and introspective protocols were both audio-recorded (Recording 2) in the same conditions as previously. Recording 2 was eventually transcribed graphemically and the dataset was coded by the researcher using the coding system as explained in section 6.2.3.

Table 156. Stages of data collection and analysis

Stage 1	L3 text reading	Recording 1
Stage 2	Analytic listening Recording 1 played back, chunks 2-3 sec long	Immediate retrospective protocol: self-correction, comments on L3 phonetic performance Recording 2
Stage 3	Stimulated recall protocols	Introspective protocol: self-reflection on L3 acquisition process Recording 2
Stage 4	Analysis	Transcription Coding Quantitative and qualitative data analysis

6.2.3. Concept operationalization and coding

The concept operationalization and the coding procedure followed, with minor adaptations, the author's proposal designed for earlier preliminary studies on metaphonological awareness (Wrembel 2013, 2015).

Based on theoretical premises, and particularly on Schmidt's (1990, 1993) noticing hypothesis, metaphonological awareness was operationalized in the present series of studies in a threefold manner as instances of:

(1) noticing, i.e. mentioning L3 phonetic features, commenting on one's own specific pronunciation problems [NOTICING],
(2) understanding, i.e. formulating phonetic rules, conscious analysis of L3 pronunciation performance [RULES],
(3) metacognition, i.e. self-reflection on L3 pronunciation and the learning process in general [COMMENTS].

Further, the conducted verbal protocols were expected to reflect different types of metaphonological awareness, following the distinction adopted from Gombert (1992):

– epilinguistic awareness, manifested by the observed instances of spontaneous correction of L3 pronunciation during the actual text reading performance [SELF-REPAIR] or during the TAP task performance [SELF-CORRECTION],
– metalinguistic phonological awareness, i.e. conscious phonetic analysis of the speech output and intentional focus on articulatory gestures reflected in the instances of noticing as well as providing rules or metacomments.

For the purpose of the objectivization of the data analysis, the present series of studies applied with minor modifications the coding system designed by the author and piloted in an exploratory study of this type (Wrembel 2015). The generated verbal protocols included two components: immediate retrospective protocols, and introspective protocols, both of which involved quantitative and qualitative data analysis.

The coding of the first part, i.e. immediate retrospective protocols, focused on the following aspects:

(1) Self-repair – ad hoc corrections of L3 phonetic features during L3 text reading performance [SELF-REPAIR],

(2) Self-correction – post hoc corrections of L3 phonetic features performed during TAP analytic listening [SELF-CORRECTION] including:
- number of instances (i.e. raw frequency),
- rate of successful vs. unsuccessful corrections.

(3) Noticing one's own problems with L3 pronunciation [NOTICING] including:
- number of instances (i.e. raw frequency),
- categorization of problems.

(4) Levels of complexity of reported metaphonological awareness [COMPLEXITY LEVEL]. This typology was adapted from Roehr's (2006: 188) work on general metalinguistic awareness and modified for the purpose of the study.
- Low complexity level – noticing and attentional focus on relevant auditory forms, i.e. phonological surface patterns [C1],
- Medium complexity level – metalinguistic description or explanation, i.e. a performed analysis of the targeted feature [C2],
- High complexity level – metalinguistic description or explanation with the use of metalanguage, i.e. the analysis is articulated employing metalinguistic categories [C3].

In the second part of the analysis, the coding of introspective protocols was carried out as follows:

(5) Reports of cross-linguistic influence related to phonetic performance in the L3, i.e., reported instances of interactions between different language systems of multilingual participants [CROSS-LINGUISTIC INFLUENCE],

(6) Statements of phonological rules [RULES],

(7) Metacognitive comments and explanations of L3 oral performance [COMMENTS], a typology adapted from Cohen (1996) as involving;
   – self-report, i.e., general statements about the learning process [REPORT],
   – self-observation, i.e. inspections of specific language behaviour [OBSERVATIONS].

The recorded oral protocols for each participant were transcribed and then subjected to the coding procedure performed by the researcher and verified three times to ensure maximum objectivity. Finally, a composite score of metaphonological awareness was calculated for each participant based on a formula proposed by the author (see section 6.3.2.4 for a detailed description of [COMPOSITE SCORE]).

6.3. Results analysis

6.3.1. Immediate retrospective protocols in TAPs

The analysis of the results of immediate retrospective verbal protocols involved instances of self-repair during L3 reading, self-correction during analytic listening, noticing one's own problems with L3 pronunciation as well as complexity levels of metaphonological awareness evidenced during reflective analysis.

In order to investigate proficiency group effects on the component measures of metaphonological awareness, each group (A, B, C, D) was further subdivided into 2 subgroups with respect to their L3 proficiency level, i.e. Groups A1, B1, C1, D1 with a lower proficiency at the A1-A2 level according to CEFR, and Groups A2, B2, C2, D2 with higher proficiency at the B1-B2 level according to CEFR. Independent samples t-tests were performed to investigate mean differences between the two proficiency subgroups and thus substantiate or disconfirm the existence of an L3 proficiency group effect on a particular aspect under investigation.

6.3.1.1. Self-corrections and self-repair

The participants in Group A showed 20 instances of self-repair during the actual reading performance in L3 French (M=0.74 per capita, SD=1.1). The nature of these ad hoc modifications concerned mostly the pronunciation of individual segments. As far as post hoc self-corrections are concerned, the total number of modifications during the analytic listening to one's recordings equalled 165, (M=6 per capita, SD=2.8) and they were correct in 63% of the cases. The mean number of self-repairs was comparable between the proficiency subgroups; however, the comparison between instances of self-corrections revealed statistically significant differences. There was a tendency for more instances of self-repair in the case of the more advanced group A2 (M=1.07, SD=1.28 vs. M=0.33, SD=0.65); however, it was not statistically significant. The higher proficiency subgroup A2 generated a higher mean of correct modifications (M=5.2, SD=2.83) than the lower proficiency subgroup A1 (M=2.17, SD=1.7, $t(25)=3$, $p<0.01$). On the other hand, the less proficient group performed on average significantly more incorrect self-corrections (M=3.75, SD=3.33) than the A2 Group (M=1.07, SD=1.03, $t(25)=2.96$, $p<0.01$).

In Group B there were 30 instances of ad hoc self-repair during the reading performance in L3 German (M=1.2, SD=1.5). In the case of self-corrections of L3 pronunciation during the analytical listening stage, the number of instances of post hoc modifications totalled 131 (M=5.24, SD=2.01). The success rate of self-corrections in Group B was 56.5%. An L3 proficiency group effect was observed with respect to self-repair and self-corrections, i.e. between the lower L3 German proficiency subgroup B1 and the higher proficiency B2. The mean number of correct modifications was significantly higher in the higher proficiency subgroup B2 (M=3.82, SD=2.33) than in the lower proficiency subgroup B1 (M=2.15, SD=1.52, $t(23)=2.14$, $p=0.04$). The opposite trend was observed for unsuccessful self-corrections, which were more numerous in the lower proficiency group (M=2.62, SD=2.14) than the more advanced group (M=1.92, SD=1.16), yet this difference did not reach statistical significance.

The first aspect of the immediate retrospective verbal protocols, i.e. self-repair during the reading performance in L3 English occurred only twice in Group C (M=0.07, SD=0.26). On the other hand, the number of post hoc self-corrections performed during the analytical listening totalled 185 with a mean of 6.38 per participant (SD=1.94). Interestingly, these modifications were correct only 29% of the time, with a prevailing percentage of incorrect attempts. In order to investigate L3 proficiency effects, independent samples t-tests were performed. While the mean number of self-repairs was comparable between the two subgroups, the more proficient C2 group generated more correct (M=2.13, SD=1.41) and incorrect self-corrections (M=5.31, SD=2.7) than the less advanced C1 group (correct modifications M=1.54, SD=1.13; incorrect modifications (M=3.54 SD=2.15); however, this tendency was not found to be statistically significant.

The participants in Group D produced only 3 instances of self-repair during the actual reading performance in L3 English (M=0.1 per capita, SD=0.33). Conversely, the number of self-corrections performed post hoc equalled 173 instances, with a mean of 6.65 (SD=1.45). Where the success rate of the performed modifications is concerned, they were correct 44% of the time, which points to the prevalence of incorrect modifications in L3 English pronunciation (56%). The conducted t-tests failed to demonstrate any proficiency group effects for the mean number of self-repairs or successful/unsuccessful modifications as their mean distribution was not found to be statistically different with respect to the higher/lower L3 proficiency subgroups (e.g. correct modifications: M=2.36, SD=1.8 in D1 vs. M=3.33, SD=2.32 in D2; incorrect modifications: M=3.64, SD=1.6 in D1 vs. M=3.8, SD=1.74 in D2).

In sum, a joint analysis of the groups points to a diverse performance in spontaneous self-repairs during L3 text reading, with Groups A and B generating a considerable number of ad hoc modifications (20 and 30 respectively) as opposed to negligible instances of self-repairs in Groups C and D. On the other hand, fairly comparable patterns were observed across all the groups with respect to self-corrections of the L3 phonetic performance. The total number of self-corrections was high in the four groups, ranging from 165 (in Group A) to 185 (in Group C), with an average mean around 6 self-corrections per person. The success rate of the performed self-corrections was on average around 50%, with the highest correctness rating in Group A (63%) and the lowest in Group C (29%). As far as the proficiency effect is concerned, the prevailing trend was for the more advanced subgroups to perform more correct modifications (Groups A, B, C) and fewer incorrect ones (Groups A, B).

6.3.1.2. Noticing of L3 pronunciation problems

Another aspect of metaphonological awareness investigated by means of the retrospective protocols concerned noticing one's own problems with L3 pronunciation.

The number of reported problems with L3 French pronunciation totalled 82 for Group A, with a mean of 3.04 instances per participant (SD=1.6). The more proficient subgroup reported a higher number of problems with L3 pronunciation (3.53, SD=1.6) than the less proficient one (M=2.42, SD=1.44); however, the difference was not significant.

In Group B the raw frequency of self-reported phonetic problems in L3 German was 89, with a mean of 3.56 instances per participant (SD=1.19). The problems were distributed relatively equally across the L3 proficiency levels as there were no significant differences across proficiency subgroups (M=3.15 vs. M=4).

Group C reported 83 instances of problems with L3 English pronunciation verbalized by the participants (M=2.86, SD=1.25). The more advanced group tended to report more problems (M=3.25, SD=1.29) than the less advanced one (M=2.38, SD=1.04), yet this trend was not statistically significant.

In Group D the participants reported 67 instances of noticed problems with L3 English pronunciation (M=2.58, SD=1.24). There was a significant difference between the means of the two L3 proficiency subgroups, with the more advanced D2 group achieving a higher mean score (M=3.13, SD=1.25) than the lower proficiency D1 group (M=1.82, SD=0.75; $t(24)=3.08$, $p<0.01$).

On the whole, the number of reported problems with L3 pronunciation clustered around the mean of 3 per participant across all groups, with the highest total number of instances in Group B (N=89) and the lowest (N=67) in Group D. A tendency to produce more comments in this respect was observed for the higher proficiency subgroups (in A, C, D); however, it only proved to be significant in Group D.

The reported problems in L3 phonetic performance were further analysed with respect to the specific categories they fell into. The most frequently noticed phonetic and phonological features concerned individual segments, both vowels and consonants, as exemplified in the following quotations, which were either quite general, e.g. *I have mostly* *problems with vowels in L3 French as consonants are similar to Polish* – Group A, *Vowels in German should be more prolonged* – Group B, *I can't differentiate between short and long vowels in English* – Group D, or more specific in nature, e.g. *Consonants /t/ /d/ are articulated in a different way, they are more aspirated in German* – Group B, *I voice /s/ to /z/ where they should not be voiced in English under German influence* – Group C, *Umlauts should be deeper, I flatten them in fast speech* – Group B, *I don't know when to read different types of /i/ in English* – Group C. Relatively less attention was devoted to suprasegmental features, with rather infrequent comments on problems with L3 word stress patterns or intonation contours, e.g. *I read everything with too flat intonation, more emphasis should be placed on the most important words in a sentence* – Group A, *In French word stress should fall always on the last syllable* – Group D.[1]

6.3.1.3. Awareness complexity levels

When the complexity of comments, corrections and reported problems with L3 pronunciation was analysed, patterns of varying complexity emerged.

In Group A, out of the total number of 258 instances, over half of all the comments (60.5%) were coded as being of a low complexity level as they mostly involved noticing a particular phonological pattern in L3 without further elaboration. One third of the comments (33.5%) were classified as consisting of the medium complexity level since they featured metalinguistic explanation, in which the participants performed conscious analysis of the targeted feature. High complexity awareness, which required the use of appropriate metalanguage in the provided explanations, was found to be the least frequent one (6%). The complexity distribution was further analysed with respect to the proficiency subgroups. While the average score was comparable for the low complexity level, significant differences were observed at the remaining levels, with higher proficiency subgroup A2 generating a greater mean for the medium complexity level (M=4.13, SD=1.68) than

[1] Quotes from the participants' introspective protocols are provided in italics. Translations from Polish into English are mine (MW).

A1 ($M=2.08$, $SD=2.07$, $t(25)=2.84$, $p<0.01$) and for the high complexity level ($M=0.93$, $SD=1.03$ vs. $M=0.08$, $SD=0.29$; $t(25)=2.76$, $p<0.05$).

For Group B, the analysis of the complexity levels of metaphonological awareness manifested through self-corrections, and reported problems with L3 German pronunciation and comments ($N=364$) pointed to the largest number of comments (56.5%) which were classified as having a low complexity level (C1), as they involved noticing a particular phonological surface pattern or providing its auditory form. Medium complexity level (C2) was evidenced in 39% of the cases, and it was characterised by the participants performing an analysis of the targeted feature accompanied by a metalinguistic description or explanation. High complexity level (C3) was found to be the least represented (7%) as it required the metalinguistic analysis to be articulated with the application of metalanguage. T-test analysis revealed interesting patterns related to proficiency group effects. While lower proficiency group B1 demonstrated a tendency for a higher number of low complexity comments, they were significantly outperformed by the more advanced group B2 ($M=8.83$, $SD=3.54$ vs. $M=2.69$, $SD=1.75$) with respect to the mid complexity level ($t(23)=5.56$, $p<0.01$) and high complexity level comments ($M=1.75$, $SD=1.6$ vs. $M= 0.38$, $SD=0.77$; $t(23)=2.76$, $p<0.05$).

As far as the complexity level analysis in Group C is concerned, out of the total number of L3 corrections, problem noticing and comments ($N=382$), the great majority (76%) was at the lowest level (C1). Mid level complexity (C2), which was characterized by some metalinguistic explanation or description, lacking at the lower level, was represented by 22% of the generated retrospective and introspective verbal protocols. The least frequent was the high complexity level (C3), reported only in 2% of cases, as it required the use of appropriate metalanguage in the provided comments. The complexity distribution was further analysed for the proficiency effect. Significant differences were observed at the low and medium complexity levels, whereas the average scores were comparable at the high complexity level. The higher L3 proficiency subgroup C2 generated a greater mean for the low complexity level ($M=11.25$, $SD=3.13$) than C1 ($M=8.46$, $SD=2.18$, $t(27)=2.7$, $p<0.01$) as well as for the medium complexity level ($M=4.06$, $SD=2.74$ vs. $M=1.54$, $SD=1.85$; $t(27)=2.82$, $p<0.01$).

As the complexity level analysis in Group D indicated, out of the total number of L3 corrections, problem noticing and comments (N=345), the majority (65.5%) was classified as low complexity level, followed by 34% of instances of awareness complexity at a medium level. Finally, only a small percentage (1.5%) of the reported instances of metaphonological awareness was classified as representing the high complexity level. An analysis of the L3 proficiency group effect on complexity level distribution failed to generate any statistically significant differences between the two proficiency subgroups as the observed means were very close to one another in D1 and D2 at all the complexity levels.

Summing up, on average nearly two thirds of the responses in the joint analysis were classified as having a low complexity characterized by more implicit forms of awareness and noticing of L3 phonological features; followed by the medium complexity level representing more explicit awareness e.g. analyzing one’s phonetic performance (less than one third of all the responses); and a small percentage of high complexity responses reflecting explicit verbalizations involving the use of appropriate metalanguage. As far as the proficiency group effect is concerned, in three out of four groups (i.e. A, B, C) the more advanced subgroups outperformed the less proficient ones in at least two of the complexity levels (usually the medium and high).

Examples of classification of the participants’ metaphonological comments into three complexity levels are presented in the following sections. In the low complexity category, in the majority of comments the participants simply echoed, i.e. repeated some words or phrases in the L3, trying to provide a modified form or expressing uncertainty as to its correct pronunciation. This level concerned noticing and attentional focus on relevant auditory forms and can be exemplified by the following statements, e.g. *I try to imitate French /r/* – Group A; *<sz> is softer in German* – Group B; *I have /s/ from Polish* – Group C.

As far as the medium complexity level [C2] is concerned, some metalinguistic description or explanation was provided or a brief analysis of the targeted feature based on the learners’ intuition or previous linguistic knowledge was performed. The following statements exemplify this category of complexity: *I pronounce French endings although they should be mute* – Group A; *The vowel in ‘schiessen’ should be more /y/ not /i/* – Group B; *I can’t articulate the dental <th>* – Group C; *German*

/r/ is harder, English /r/ is softer – Group C; *I can't distinguish the length of the vowels, I reduce them all to short vowels* – Group C; *In <-ing> endings final <g> should be mute* – Group D.

In the case of the high complexity level [C3], metalinguistic descriptions or explanations were provided with the use of appropriate metalanguage, as exemplified in the selected quotations: *Influence from L2 English can be observed in too aspirated plosives* – Group A; *There should be more vowel reduction* – Group A; *French has final devoicing like Polish* – Group A; *The initial sound should be a schwa not a diphthong* – Group B; *I am used to ultimate word stress in French, in English it is not so, you have to learn a word with its stress pattern* – Group D.

6.3.2. Introspective verbal protocols

The analysis of introspective verbal protocols included reported cross-linguistic influence (CLI), the formulations of phonological rules as well as metacognitive comments and observations on the L3 oral performance. The statements referred to salient features of L3, L2 and L1 sound systems, featuring conscious phonetic analysis and the use of metalanguage.

6.3.2.1. Reported cross-linguistic influence

The first category of metacomments generated in the introspective verbal protocol concerned the phenomenon of cross-linguistic influence. The question under investigation was which of the existing language systems a learner transfers from in the process of the acquisition of third language phonology. The reports reflected perceived interactions between different language systems of the multilingual participants, and involved declared sources of phonological transfer, as well as its directionality and strength. In the self-reports on CLI, the respondents identified also the reasons and factors guiding this phenomenon.

Three categories of responses were provided that identified:

- L1 transfer as the basic constraint on articulation
- L2 transfer, i.e. a tendency to activate L2 in L3 production
- Combined L1/L2 transfer, i.e. an impact of all the previously acquired languages.

The following sections present the quantitative and qualitative analysis of the reported instances of CLI.

As far as the quantitative analysis of reported CLI is concerned, in Group A 68 instances of metacomments were classified as reporting cross-linguistic influence (M=2.52, SD=0.98). The comparison between different proficiency subgroups did not yield any statistically significant results, with the lower proficiency group scoring on average M=2.33 (SD=0.98) and the more advanced group achieving similar results (M=2.67, SD=0.98).

The overall number of reports of cross-linguistic influence in Group B totalled 93 (M=3.72, SD=2.01). Statistically significant differences were found between the higher proficiency subgroup (M=4.83, SD=1.8), which scored better and reported more instances of CLI, and the less advanced group (M=2.69, SD=1.65; $t(23)=3.1$, $p<0.01$).

In Group C the number of instances of reported CLI amounted to 86 (M=2.97, SD=1.64). Similarly to Group B, the scores of the less proficient subgroup were significantly lower (M=2.15, SD=0.99) than those of the higher proficiency subgroup (M=3.63, SD=1.78).

For Group D, 80 instances of comments on CLI were reported (M=3, SD=1.13). No L3 proficiency group effect was attested as the mean score for both proficiency subgroups were comparable (lower proficiency: M=3.36, SD=1.03 vs. higher proficiency: M=2.87, SD=1.19).

All in all, the reported cross-linguistic influence was at a fairly comparable level across all the groups, ranging from 68 to 93 instances per group. The highest mean value was observed in Group B with L3 German (M=3.72), whereas the lowest in Group A with L3 French (M=2.52), yet the differences between groups were not substantial. Only in two groups (B, C) were L3 proficiency group effects observed, thus pointing to higher rates of reported CLI in more advanced groups, as could have been expected.

Examples of metacomments on cross-linguistic influence as reported by the participants are presented in the following sections as part of a qualitative analysis. The vast majority reported that their second language (L2) prevails as the source of transfer in the phonological acquisition of a third language and as an external supplier language for articulatory patterns in L3 (48% of L2 English-to-L3 French CLI in Group A; 62.5% of L2 English-to-L3 German CLI in Group B; as much as 86% of L2

German-to-L3 English in Group C, and 50% of L2 French-to-L3 English in Group D). This corroborated, to a large extent, the assumption of the foreign language effect, or 'L2 status', commonly acknowledged in the recent L3 literature, according to which the second language constitutes a prevailing source of cross-linguistic influence in the acquisition of a subsequent language, and there is a tendency to activate the L2 in L3 production (cf. Cenoz, 2001; De Angelis, 2007; Hammarberg and Hammarberg, 2005). Interestingly, if we allow for typological proximity between the language pairings and compare L2/L3 mirror combinations, we can observe an additional effect of typology, as the strength of the reported L2 influence was found to be considerably higher in the typologically more closely related English and German (86% in Group C, 62.5% in Group B) than in the more distant English and French pairings (48% in Group A, 50% in Group D).

A significantly smaller percentage of the participants indicated their L1 as the main source of cross-linguistic phonological influence in the acquisition of their respective third language (22% of L1 Polish-to-L3 French CLI in Group A; 12.5% of L1 Polish-to-L3 German CLI in Group B; 10% of L1 Polish-to-L3 English in Group C; and 23% of L1 Polish-to-L3 English in Group D). This finding provided only weak support for the traditionally held belief that the articulatory motor routines established in the mother tongue determine the acquisition of any subsequent language systems, be it the second or third language (Ringbom, 1987). Notably, this influence was reported as weaker in English and German L2/L3 mirror pairings (12.5% and 10%), and somewhat stronger in the more distant English/French L2/L3 groupings (22% and 23%). It may be concluded that the reported L1 transfer was modified to some extent by the typological proximity between the remaining foreign language systems (i.e. the stronger the typological closeness between the non-native languages, the weaker the reported impact of the native language on the L3).

There were also indications of a combined cross-linguistic influence, i.e. in which multiple sources of interference were recognized, mostly involving both the L1 and L2: (30% of combined CLI in Group A, 25% in Group B, 4% in Group C and 27% in Group D). Interestingly, the percentage of reported combined influence from both native and non-native languages on L3 was fairly comparable in three out of the four groups. This category of responses lends support to De Angelis' (2005) proposal of a combined cross-linguistic influence that was, in turn, corroborated by the research results (cf. Benrabah 1991, Blank and Zimmer 2009, Hammarberg and Hammarberg 2005, Wrembel 2010, 2012, Wunder 2010).

The participants made attempts to account for the reported sources of CLI by indicating various factors that determine this phenomenon. Specifically, they frequently pointed to different routes of acquisition differentiating between the naturalistic acquisition of the first language vs. the formal learning of foreign languages, be it L2, L3 or Ln (e.g. *Separate categories are created for the native and foreign languages, that is why there are more interactions between L2 and L3; Both L2 and L3 were learnt as foreign languages, that's why there is more mixing between them* – Group A; *I treat foreign languages as a separate category, I don't identify so much with them as with my native language, they are not so firmly embedded* – Group C). Some of the comments related directly to cognitive associations between non-native languages and the so called 'switch to a foreign language mode' when performing in any foreign language, which is one of the cognitive explanations provided for the L2 status effect (cf. Cohen 1995), e.g. *I automatically switch into a foreign language* – Group B and C; *I have an automatic influence from my first foreign language even though German phonetics seems different from English phonetics* – Group C. Some of the reports also confirmed Fernandes-Boëchat's (2007) assumption that the process of learning a new foreign language is linked involuntarily to the preceding foreign language learning experience.

On the other hand, many reports claimed that due to its different status the native language does not act as the main source of CLI for a third language (e.g. *Polish is my inner language, L2 and L3 are additional 'unnatural' languages, that is why they do not mix, I differentiate native language from foreign languages therefore there is no influence from L1* – Group A). Although less frequent, there were also some opinions voiced in favour of L1 Polish as the main source of cross-linguistic influence. The arguments provided pointed to the neuro-motor routines formed in the process of the first language acquisition that condition the pronunciation of any subsequently learnt languages (e.g. *I automatically resort to Polish, Phonetic schemes that are transferred from L1 are unconscious* – Group D).

The most frequent explanation for the prevalence of the second language as a source of CLI provided by the participants was the amount of exposure to the second language as well as the recency of use (e.g. *Because of the constant focus on English pronunciation I transfer it into L3 French* – Group A; *English is always on my mind* – Group A, B; *I have a lot of constant exposure to German and German music* – Group C; *French is always at the back of my mind* – Group D). Some participants also claimed that the L2 effect on L3 is determined by the type of instruction and instructional setting (e.g. *Because it's another foreign language that was learnt in a similar school setting* – Group C; *Due to phonetic training in L2 English that I received I tend to transfer phonetic experience from English* – Group B).

Moreover, the chronology of acquisition was also reported to condition the source and directionality of CLI, i.e. the language acquired later was claimed to be influenced by the preceding one, but not vice versa (e.g. *English is my first foreign language, thus it influences my German* – Group B, *English was the first foreign language, it is so strongly embedded that I subconsciously resort to it* – Group A; *German was my first foreign language, thus I associate everything else with German* – Group C).

As far as the stages of L3 acquisition are concerned, they seem to have some potential bearing on the selection of the source of transfer. Some reports stressed that the influence from the L2 was particularly strong in the early stages of the acquisition of the third language and that it diminished with time (e.g. *L2 influence especially noticeable at the beginning of L3 acquisition* – Group A) which is consistent with the literature (cf. Hammarberg and Hammarberg 2005, Wrembel 2010).

Interesting observations concerned a conscious strategy acknowledged by some participants who admitted that they try to suppress their L1 Polish accent consciously (e.g. *I prefer to sound more English than Polish in my L3 German, It's better to sound English-accented* – Group B). Such a strategy was first identified in the L3 literature by Hammarberg and Hammarberg (1993, 2005) as a foreign language coping strategy in which conscious suppression of the native language was accompanied by the reactivation of the second language strategies. Moreover, De Angelis (2005, 2007) referred to it as the reaction to two opposing trends, namely, the perception of correctness vs. the association of foreigness.

Typological proximity between languages has frequently been acknowledged as a factor conditioning the source of transfer (e.g. *German and English are typologically related, they come from the same language family* – Group B; *Both German and English are Germanic languages, so we look for common features between them* – Group C; *Polish is a Slavic language, typologically distant from English, that's why there is no influence from it* – Group C). Some of the comments seem contradictory, yet they relate to a subcategory of typological proximity, namely, psychotypology – which is more a subjective than systemic perception of typological closeness or the distance between language systems, as illustrated in the following examples (e.g. *Polish and French are typologically similar in terms of pronunciation; I group French and English together as they are quite similar in terms of phonology and lexis* – Group A; *Polish and French are typologically closer than English and French* – Group D).

Some of the reports specified particular areas of phonology which were affected by cross-linguistic influence (e.g. *I transfer all foreign vowels and consonants into my L3* – Group B; *I catch myself trying to pronounce many vowel and consonant sounds in French in an English manner* – Group A; *The short/long vowel distinction is transferred from German to my English* – Group C; *I can hear Polish influence on my English in those hard consonants* – Group D). Others differentiate between sources of potential transfer with respect to the degree of consciousness and control during their L3 oral performance (e.g. *When I focus on L3 pronunciation I try to make it sound foreign and thus transfer from L2 English, when I don't focus and speak naturally, I have more transfer from L1 Polis*h – Group B; *I have more influence from French in my English when reading* – Group D).

More infrequent comments relate to a perceived separation between particular language systems that prevents mixing and apparently blocks CLI (e.g. *I try not to mix German and English, I have separate 'drawers' in my mind for these languages* – Group B; *There is no influence from L2 German on my L3 English, I am able to keep these two foreign languages separate* – Group C).

As far as the different directionality of cross-linguistic influence is concerned, some participants acknowledged that their dominant L2 influences not only the L3 but also their native language (L1), a trend which is referred to in the literature as 'regressive transfer'. Furthermore, there were instances when the direction of transfer was subject to change in the course of language learning due to a changing dominance and a resulting shift in the L2/L3 status (e.g. *It used to be different when English was my dominant foreign language than it was influenced by L1 Polish, now when French is my stronger foreign language, it influences my English sometimes* – Group D). Finally, some explanations were provided indicating what conditions the direction of the CLI (e.g. *The directionality of the influence (L2 - L3) depends on how well you know a foreign language* – Group C) in the sense that it is the dominant foreign language that influences the weaker one rather than the reverse. Some individual participants declared instances of transfer onto L3 from other languages they knew such as Latin or L4 French or Spanish.

Particularly noteworthy is that the majority of the factors acknowledged by the participants as conditioning the process of cross-linguistic influence, based on their intuitions and introspections, reflected the exact set of determiners that are identified in the third language acquisition literature (cf. Cenoz, 2001; De Angelis, 2007; Hammarberg, 2009). Recapitulating, it would be interesting to juxtapose the self-reported data on the sources of transfer with the participants' actual performance in the L3. Such a comparison between the results of introspective verbal protocols and performed foreign accentedness ratings, as well as VOT measurements with respect to the sources and directionality of cross-linguistic influence, is presented in Chapter Seven, section 7.1.

6.3.2.2. Formulations of phonological rules

The following category of introspective comments generated in the verbal protocols involved the formulations of phonological rules by the participants. These attempts were usually characterised by conscious phonetic analysis and the application of appropriate metalanguage. They referred to the features of one or several of the sound systems of the multilingual participants that they considered particularly salient. In the sections below, both the qualitative and quantitative analysis of phonological rules will be provided.

As far as the quantitative analysis is concerned, in Group A the number of provided phonological rules accounting for the participants' performance in L3 French equalled 30 instances (M=1.1 per capita, SD=1.09). No significant L3 proficiency group effect was found, although the higher proficiency subgroup tended to provide on average more phonological rules (M=1.4, SD=1.18) than the less proficient one (M=0.75, SD=0.87).

In Group B, the participants generated 33 statements that could be classified as phonological rules (M=1.32, SD=1.28). More instances of rules were provided by the higher proficiency subgroup (M=2, SD=1.28) than the less advanced one (M=0.69, SD=0.95) and the difference was found to be statistically significant ($t(23)=2.92$, $p<0.01$).

A similar pattern was attested in Group C, with 29 instances of stated phonological rules (M=1, SD=1). Likewise, the comparison between L3 proficiency subgroups pointed to a statistically significant difference, as the more advanced group outperformed the less proficient one (M=1.38, SD=0.96 vs. M=0.54, SD=0.88; $t(27)=2.43$, $p<0.05$).

As for Group D, the number of generated phonological rules totalled 23 (M=0.88, SD=0.91). No L3 proficiency group effect was attested in this case, as the mean scores were comparable for the less proficient (M=1, SD=0.63) and more advanced subgroups (M=0.8, SD=1.08).

Summing up, the number of statements of phonological rules was distributed rather equally across the groups, with the minimum of 23 instances in Group D, and the remaining groups (B, C, D) with scores clustered around 30. There were no significant differences between the mean scores, the lowest being in the case of Group D (M=0.88) and the highest for Group B (M=1.32). An L3 proficiency group effect on the number of stated rules was evidenced in three out of four groups; in Groups B and C it reached statistical significance, in Group A it was a visible trend that more proficient learners were able to provide more instances of phonological rules.

The examples of phonological rules generated by the participants are presented in the subsequent sections as part of a qualitative analysis of introspective metacomments. The majority of the provided comments in the form of rules referred to the segmental features, both to the vowels and consonants of the respective third languages (i.e. English, French and German). Several of the rules concerned the vocalic systems, e.g. the presence or lack of distinction between tense and lax vowels, e.g. *Vowel length is distinctive in German; Vowels in German should be more prolonged* – Group B. Moreover, cross-linguistic differences between particular languages in this respect were acknowledged (e.g. *Nasal vowels in French are the same as in Polish* – Group A, *German /a/ is similar to English ash* – Group B). The consonantal repertoire was equally attended to and the following statements can illustrate the provided rules: *Whenever /s/ is in an intervocalic position is should be voiced to /z/, /r/ should be more uvular; Consonants /t/ /d/ are articulated in a different way, they are more aspirated in German* – Group B, *In <-ing> endings final <g> should be mute* – Group D, */h/ dropping is a basic rule in French* – Group A.

The participants commented also on universal processes comparing their application cross-linguistically (e.g. *Initial and final devoicing exists in German as it does in Polish* – Group B, *French has final devoicing like Polish* – Group A). Some formulations of rules involved articulatory descriptions, e.g. *To pronounce <th> I touch my upper teeth with my tongue* – Group C. Significantly fewer rules concerned suprasegmental features such as word stress or intonation patterns (e.g. *In French, word stress falls always on the last syllable, I am used to ultimate word stress in French, in English it is not so, you have to learn a word with its stress pattern* – Group D; *Intonation should be falling at the end of a sentence* – Group A) or processes of casual speech (e.g. *There should be more vowel reduction* – Group A).

Furthermore, the participants' comments formulated as rules, referred sometimes also to the spelling and sound correspondence (e.g. *The ending <ent> in 3rd person plural is mute in French; You don't read the ending in a verb in imparfait* – Group A). Other comments in the form of rules concerning consonants focused on the manner of the articulation of particular foreign sounds, like umlaut or /r/ in German (e.g. *In German there are three types of /r/, I can produce them but not always in appropriate positions* – Group B).

6.3.2.3. Metacognitive comments: Self-report

Instances of self-report that involved general statements about the process of learning L3 pronunciation voiced by the participants as part of the introspective oral protocols were subdivided by the present author into several categories that reflected different aspects of metaphonological awareness such as conscious control, cognitive processing, frame of mind, learning goals and priorities, general evaluation and multilingual advantage.

As far as the first aspect is concerned, the participants revealed quite a considerable degree of attempted control over their L3 oral performance. However, the majority of comments which evidenced a failure in exerting conscious control over the process of pronunciation in the third language pointed to the difficulties and potential costs involved (e.g. *It is difficult to control L3 pronunciation consciously* – Group A; *When I try to monitor pronunciation, it affects the coherence of speech* – Group A) as well as specified the reasons (e.g. *I focus on content rather than the form of speech* – Group A; *I can't fully control my pronunciation in L3 German, I don't have enough phonological knowledge* – Group B; *I can't control my pronunciation, especially under stress* – Group D). On the other hand, some participants stated that they were able to consciously pay more attention to pronunciation performance in L3 under certain circumstances (e.g. *I control more my pronunciation when I speak with native speakers of German* – Group B).

Self-reports of cognitive processing concerned mainly code-switching between different language systems, which was generally acknowledged as problematic for the multilingual participants as it required special articulatory adjustment or re-setting (e.g. *It's difficult to switch into English after speaking French* – Group D; *When I code-switch between languages, there is a lag in terms of phonetics, a short transition period before the articulators adjust to another setting* – Group B). There was an attempt to provide a neurobiological explanation for this process (e.g. *I find it difficult to switch from one foreign language to another as if one part of the brain was responsible for it* – Group D). Moreover, the participants reported a certain degree of cognitive confusion caused by the co-existence of several language systems (e.g. *Foreign languages get mixed up in my mind* – Group D; *I automatically transfer German phonetic rules into English* – Group C) and pointed to instances of a detrimental effect of one language over another (e.g. *The more I learn French, the more my English undergoes attrition* – Group B). On the other hand, there were indications of the cognitive facilitation of linguistic intuition for additional language learning *(e.g. Linguistic intuition helps me to deduce certain things* – Group B).

Another aspect of metaphonological self-report was manifested in multilinguals assuming a novel frame of mind caused by increased metalinguistic awareness as well as perceptual sensitivity (e.g. *When you've learnt the phonetics of a foreign language, you think about this language in a completely different way* – Group A; *Once you get to know one language, you can perceive much more in another language* – Group A). The participants pointed also to the factors conditioning this cognitive facilitation in additional language learning, including a broadened phonetic repertoire as well as the recognition of typological proximity (e.g. *An awareness that there is a different repertoire of vowels and phonetic features helps in L3 acquisition* – Group A; *There is facilitation when I learn languages from the same language family* – Group A).

Pronunciation attitudes and goals featured as yet another aspect of self-report evidenced in the introspective protocols. The participants frequently expressed their positive attitudes and emotional reactions towards the target model of the third language, though it concerned mostly L3 English (e.g. *I like the British accent, I'd like to speak with a British accent, it's very melodious* – Group D; *I'd like to have a more British accent* – Group C; *I like the American accent, I try to Americanise my pronunciation but what comes out sounds Polish* – Group D). Interestingly, several comments reflected the learners' desire to speak with a native-like accent in their third language which contrasts with common contemporary trends promoting comfortable intelligibility as a sufficient language learning priority (e.g. *I would like to speak as British people do* – Group D; *When I speak I try to pretend I am a British person* – Group D).

Finally, self-reported statements about the process of learning L3 pronunciation involved also general evaluation and recommendations. Several participants acknowledged that pronunciation was an important aspect of foreign language learning for them and provided some explanatory grounds for it (e.g. *I am aware that pronunciation is important* – Group A; *Pronunciation is important, I pay attention to it, I'd like to be well perceived* – Group B) as well as pedagogical implications advocating the need for pronunciation teaching in the school curricula (e.g. *It is wrong that pronunciation is not taught at schools, even at an advanced level, because poor pronunciation can impede communication* – Group C). However, opposing views were also expressed reflecting the alleged superiority of intelligibility over native-like accent in international communication (e.g.

Everybody speaks English in their own way, nobody pays too much attention to accent as long as it is comprehensible – Group D).

Some of the participants' introspective comments concerned the so called multilingual advantage, i.e. whether they are better equipped linguistically and/or cognitively than monolinguals for the task of continued phonological acquisition. They pointed out that due to their previously acquired first foreign language (L2) they have gained language-learning experience and strategies (e.g. *Once you learn a foreign language, other languages are easier, all European languages are similar, you just change some rules, but there is the same schema* – Group A). Further, they indicated that they have now a broadened phonetic repertoire that facilitates subsequent language learning (e.g. *I am aware that languages have different repertoires of vowels, consonants and phonetic features and I now pay more attention to it* – Group B). An important aspect that was frequently commented upon was increased metalinguistic awareness and cognitive flexibility (e.g. *The more I learn languages, the more I see interconnections between them, the easier it is to learn rules, the more my language learning 'apparatus' is developed* – Group A). Finally, the multilingual participants expressed their recognition of enhanced perceptual sensitivity being reflected in their greater ease of discriminating non-native sounds (e.g. *Once you get to know one language, you can perceive much more in another language* – Group A; *I can now see differences between different types of /r/, open and close vowels, I can hear that now, earlier I did not pay attention* – Group D). Only rarely did the participants voice any scepticism related to there being an apparent multilingual advantage (e.g. *Other languages can have a positive but also negative effect, you have to think about too many factors* – Group C).

6.3.2.4. Metacognitive comments: Self-observations

Among the metacognitive comments elicited through the introspective verbal protocols two major categories were distinguished including self-report, i.e. the participants' general statements about the process of learning a third language (discussed in 6.3.2.3) and self-observation, i.e. comments on and explanations of specific language behaviour related to L3 pronunciation, which will now be presented in this section. The reported observations concerned mostly articulatory problems in the L3, instances of cross-linguistic comparisons, statements of one's phonetic strengths and weakness, remarks on specific learning strategies applied as well as general reflections on L3 pronunciation and the perception of one's accent. All the aforementioned aspects will be briefly discussed and illustrated with appropriate quotations from the participants' introspective verbal protocols.

The participants made several self-observations concerning concrete articulatory problems and a raised awareness of articulatory posture and settings (e.g. *I have more awareness of articulatory gestures* – Group A; *I can't remember how to shape my lips to articulate the sound properly* – Group B; *I have problems with opening my mouth, if I opened my lips more broadly, the sounds would be clearer* – Group C; *I feel as if I had a lump in my mouth when I speak English under the influence of German* – Group C). On the other hand, specific perceptual problems in L3 found also their reflection in the verbal protocols, thus evidencing the participants' awareness of non-native phonetic distinctions they have not mastered yet (e.g. *I can't differentiate between short and long vowels in English* – Group D; *I can't hear a difference between long and short /o/ and /i/* – Group D).

Interestingly, the self-observations included also attempts at analysing one's strengths and weaknesses as far as L3 pronunciation is concerned, thus evidencing the participants' raised awareness in this area (e.g. *I focus more on consonants while my vowels in L3 French are weaker* – Group A; *I have mostly problems with vowels in L3 French as consonants are similar to Polish* – Group A; *I can't distinguish the length of the vowels, I reduce them all to short vowels* – Group C; *I can't articulate the dental <th> in English* – Group C).

An aspect of cross-linguistic awareness characteristic for multilingual speakers, as noted by Jessner (2006), found its reflection in numerous comparisons between the various language systems at their disposal uttered as part of introspective protocols in all the groups. These cross-linguistic comparisons seemed to present a very frequent reflection of multilingual metaphonological awareness. They ranged from more general remarks (e.g. *There is a difference between English and German, English is such a soft language, German is hard* – Group C; *In French you don't pronounce endings, it gets transferred to my English* – Group D; *Because of the German influence I pronounce English in a harder manner* – Group C; *When some words are spelt similarly in French and English I* *transfer L3 English pronunciation to L3 French* – Group A) to very specific observations (e.g. *In German word endings I make nasal plosion as in English* – Group B; *I have a hissing /s/ from Polish* – Group C; *German and English /r/ in the word final position is different; however, for me this difference is blurred* – Group C; *Influence from L2 English can be observed in too aspirated plosives in my French* – Group A; *I voice /s/ to /z/ where they should not be voiced in English under German influence* – Group C; *I tend to forget to pronounce the initial /h/ in English, after a long period of use of French* – Group D; *In French we have elision, it influences my English* – Group D).

As far as the perception of accentedness in L3 is concerned, several participants indicated that their L3 oral performance tended to be L2-accented rather than L1-accented (e.g. *I can hear French-accented English in other students* – Group A; *I have a German accent in English* – Group C; *I don't have such a hard Polish accent in my English* – Group B).

The participants pointed to several strategies that they used consciously in order to address phonetic difficulties in their L3 usually by resorting to some features or processes from their native tongue (e.g. *When I don't know how to pronounce a vowel I replace it with Polish* – Group A; *I replace German vowels with Polish equivalents* – Group B) or the previously acquired foreign language (e.g. *I try not to pronounce Polish /r/, instead I try to replace German /r/ with some kind of English sound* – Group B). Some observations concerned also the facilitatory effect of other foreign languages (e.g. *It's easier to learn English and Spanish thanks to French* – Group D).

The final category of self-observations involved expressions of uncertainty or lack of adequate knowledge (e.g. *I can hear something is wrong with my L3 French pronunciation but I can't correct it* – Group A; *I don't know if the velar nasal should be followed by /g/ or not* – Group B; *I have problems with pronouncing English words of French origin* – Group D; *Phonetic rules of French and English get mixed up* – Group D) as well as the participants' own critical reflexivity on what should be improved in their L3 pronunciation (e.g. *I should finish this sentence with a higher intonation contour* – Group B; *I read everything with too flat intonation, more emphasis should be placed on the most important words in a sentence* – Group A; *I prolong vowels in French at the end of words to gain fluency* – Group A).

6.3.3. MPhA composite score

The following formula was proposed by the present author to account holistically for metaphonological awareness (MPhA) as a composite score of various parameters that were categorized in the oral protocol analysis.

A*2+B-C*0.5+D+E+F*2+G*3+H+I*2-J*0.5+K=MPhA composite score

- A – Pronunciation ad hoc self-repair (x 2)
- B – Pronunciation post hoc self-correction (x 1)
- C – Incorrect correction (x -0.5)
- D – Noticing pronunciation problems (x 1)
- E – Low complexity level (x 1)
- F – Medium complexity level (x 2)
- G – High complexity level (x 3)
- H – Reported cross-linguistic influence (x 1)
- I – Formulation of phonological rules (x 2)
- J – Expressions of uncertainty (x -0.5)
- K – Metacognitive comments (x 1)

To calculate the MPhA composite score, for each participant the number of observed instances in each category was multiplied by an indicator of its relative importance (see figures in parentheses) and the total sum was arrived at. The indicator of importance was provided in an attempt to weigh the relative contribution of particular parameters to the composite measure of metaphonological awareness. The basic score of 1 point was assigned to several parameters including L3 pronunciation self-correction, noticing problems in L3 pronunciation, reported cross-linguistic influence, metacognitive comments and instances of low complexity level awareness. A higher measure of 2 points was allocated to parameters that required more advanced manifestations of MPhA such as formulations of phonological rules, medium complexity level metacomments or more automatized self-repair in L3 pronunciation. The highest weighting of 3 points was assigned only in the case of the high level of complexity, i.e. when metalinguistic comments were accompanied by an explanation and appropriate use of the metalanguage. Negative values of -0.5 point were allocated to two subcomponents including expressions of uncertainty and explicit lack of metalinguistic knowledge as well as instances of self-corrections of L3 pronunciation that were not correct. The MPhA composite score, stemming from the application of the formula, was calculated for each individual participant. Further, on the basis of the composite scores, three levels of meatphonological awareness were distinguished including: low awareness (below 20 points), medium awareness (between 20-40 points) and high awareness (above 40 points). Finally, mean results for particular groups (A, B, C, D) as well as proficiency subgroups within these groups were calculated and percentages of participants assigned to specified MPhA levels were assessed. The findings for individual groups will be presented in the sections that follow and summarized in Table 157.

The mean composite score of MPhA for Group A was 28.3 (SD=12), ranging from 6 to 54.5 points, with 26% of the participants falling into the low awareness category (<20 points), 59% into medium awareness (20-40 points) and 15% in the high awareness group (>40 points). As far as the proficiency group effect on the MPhA composite score is concerned, the more advanced subgroup A2 performed significantly better (M=36, SD=7.78; $t(25)=5.48$, $p<0.01$) than the lower proficiency subgroup A1 (M=18.71, SD=8.59).

In the case of Group B, the mean composite score equalled 39.48 (SD=17.59), within a range of a minimum 15 to a maximum 73.5 points. The category distribution was as follows: 20% of the participants were classified as having low metaphonological awareness (<20 points), 24% of the participants as belonging to the medium awareness category (20-40 points), and the majority, i.e. 56% representing the high awareness group (>40 points). Statistically significant differences in the MPhA composite score were found between L3 proficiency subgroups; with the higher L3 German proficiency B2 group (M=51.88, SD=12.62) outperforming the less advanced B1 group (M=28.04, SD=13.32; $t(23)=4.58$, $p<0.01$).

For Group C, the MPhA composite score amounted on average to 26.31 (SD=10.52), ranging from 9 to 43 points. A detailed analysis shows that 38% of the participants fell into the low metaphonological awareness group (<20 points), a majority, i.e. 48% to the medium level awareness, and only 14% were classified as highly aware participants. The MPhA composite scores differed significantly with respect to L3 English proficiency subgroups with the more advanced C2 subgroup (M=31.56, SD=9.64) performing much better than the lower proficiency C1 subgroup (M=19.85, SD=7.75; t(27)=3.54, p<0.01).

As for Group D, the metaphonological composite score had a mean of 28.31 (SD=11.85) and a minimum vs. maximum range of 6 to 54.5. In the degree of awareness category distribution 26% of the participants scored in the low awareness range (<20 points), a great majority, i.e. 59% were classified at the medium awareness level (20-40 points) and 15% of the participants ranked as highly metaphonologically aware (>40 points). The higher L3 English proficiency subgroup D2 demonstrated on average better MPhA composite scores (M=32.43, SD=12.36) than the less advanced D1 subgroup (M=28.5, SD=8.07); however, this difference was not found to be statistically significant.

An across-group comparison of average results for all the parameters as well as the composite score of metaphonological awareness is presented in Table 157. Significant proficiency effects within subgroups are marked with an asterisk.

Table 157. Metaphonological awareness scores for all parameters

MPhA	Group A	Group B	Group C	Group D
Self-repair	M=0.74 (1.1)	M=1.2 (1.5)	M=0.07 (0.3)	M=0.1 (0.3)
Self-correction	M=6.0 (2.8)*	M=5.2 (2.0)*	M=6.4 (1.9)	M=6.6 (1.4)
Pronunciation problems noticing	M=3.0 (1.6)	M=3.6 (1.2)	M=2.9 (1.2)	M=2.6 (1.2)*
Comments complexity levels	C1 60.5% C2 33.5%* C3 6%*	C1 56.5% C2 39%* C3 7%*	C1 76%* C2 22%* C3 2%	C1 65.5% C2 34% C3 1.5%
Reported CLI	M= 2.5 (0.9)	M=3.7 (2.0)*	M=2.9 (1.6)	M=3 (1.1)
Sources of CLI	L1 22% L2 48% Mix 30%	L1 12.5% L2 62.5% Mix 25%	L1 10% L2 86% Mix 4%	L1 23% L2 50% Mix 27%
Formulation of phonological rules	M=1.1 (1.1)	M=1.3 (1.3)*	M=1.0 (1.0)*	M=0.9 (0.9)
MPhA composite score	M=28.3 (12)*	M=39.5 (17.6)*	M=26.3 (10.5)*	M=28.3 (11.9)
Participants' awareness levels	Low 26% Mid 59% High 15%	Low 20% Mid 24% High 56%	Low 38% Mid 48% High 14%	Low 26% Mid 59% High 15%

6.4. Correlational analyses

The following sections aim to investigate interrelations between metaphonological awareness, quantified by means of the MPhA composite score, and (1) separate component parameters of metaphonological awareness, (2) accentedness ratings from Study 1, (3) participants' variables.

6.4.1. Correleations between the MPhA components and composite score

Correlations between particular component parameters of the metaphonological awareness and the composite score were investigated by means of a Spearman correlation test for all language groupings treated jointly. Several statistically significant correlations were found to hold between the composite score and nearly all the component parameters (see Table 158). Moderate correlations were observed between the composite score and such variables as noticing pronunciation problems in L3 (R_S=0.41); and instances of ad-hoc self-repair in the L3 reading performance (R_S=0.33). The variables that correlated highly with the awareness composite score included the number of correct corrections (R_S=0.55); the formulations of phonological rules (R_S=0.55); the instances of reported CLI (R_S=0.58); and the C3 level of awareness reflected in the metacomments (R_S=0.61). A very high correlation was found to hold between the composite score and the C2 level of awareness (R_S=0.80). The only component parameters that did not exhibit significant correlations included instances of incorrect modifications, the C1 level of awareness and the expressions of uncertainty or lack of knowledge. On the whole, the results lend internal validity to the composite measure of metaphonological awareness based on the proposed formula ascribing different weight to particular components of MPhA, based on their relative importance.

Table 158. Between components correlations – a joint analysis

Pairs of variables	N	R_s	t(n-2)	p
Self-repair & Composite score	107	**0.33**	3.53	**0.000612**
Correct correction & Composite score	107	**0.55**	6.72	**0.000000**
Incorrect correction & Composite score	107	-0.14	-1.47	0.144060
Noticing problems & Composite score	107	**0.41**	4.59	**0.000012**

C1 level & Composite score	107	0.18	1.84	0.069061
C2 level & Composite score	107	**0.80**	13.74	**0.000000**
C3 level & Composite score	107	**0.61**	7.80	**0.000000**
Reported CLI & Composite score	107	**0.58**	7.26	**0.000000**
Formulating rules & Composite score	107	**0.55**	6.74	**0.000000**
Uncertainty & Composite score	107	-0.07	-0.73	0.465825
Metacomments & Composite score	107	**0.60**	7.62	**0.000000**

6.4.2. Correlations between accentedness ratings and metaphonological awareness

Spearman rank correlation tests were performed to investigate correlations between accentedness ratings (Study 1) and metaphonological awareness scores (Study 3) in all the respective groups.

In Group A, all the aspects of pronunciation performance in L3 French evaluated by means of ratings, correlated moderately with the participants' composite score for metaphonological awareness, i.e. metaphonological awareness and accent R_S=0.46, metaphonological awareness and comprehensibility R_S=0.45, and metaphonological awareness and pronunciation correctness R_S=0.38, $p<0.05$ (see Table 159). It follows that the higher the composite score of the metaphonological awareness test, the higher the participants scored for their L3 pronunciation performance in terms of sounding less foreign accented, more comprehensible and correct.

Table 159. Spearman rank correlations between accentedness ratings and MPhA for Group A with L3 French

Pairs of variables	N	Rs	t(n-2)	p
Foreign accent & MPhA score	27	**0.46**	2.60	**0.015405**
Comprehensibility & MPhA score	27	**0.45**	2.51	**0.019108**
Correctness & MPhA score	27	**0.38**	2.07	**0.048635**

In Group B no significant Spearman rank correlations were found for any aspect of pronunciation performance in L3 German evaluated by means of accentedness ratings and the participants' composite score for metaphonological awareness (see Table 160).

Table 160. Spearman rank correlations between accentedness ratings and MPhA for Group B with L3 German

Pairs of variables	N	R_S	t(n-2)	p
Foreign accent & MPhA score	25	0.12	0.57	0.571143
Comprehensibility & MPhA score	25	0.00	0.01	0.990523
Correctness & MPhA score	25	0.00	0.01	0.989804

Similar results were generated for the remaining groups as there were no significant correlations between any of the aspects of accentedness ratings including foreign accent, comprehensibility, correctness and, on the other hand, the composite score of metaphonological awareness in Groups C and D with L3 English (see Tables 161 and 162).

Table 161. Spearman rank correlations between accentedness ratings and MPhA for Group C with L3 English/L2 German

Pairs of variables	N	R_S	t(n-2)	p
Foreign accent & MPhA score	16	-0.09	-0.33	0.744858
Comprehensibility & MPhA score	16	-0.04	-0.14	0.887705
Correctness & MPhA score	16	-0.07	-0.25	0.807340

Table 162. Spearman rank correlations between accentedness ratings and MPhA for Group D with L3 English/L2 French

Pairs of variables	N	R_S	t(n-2)	p
Foreign accent & MPhA score	15	-0.17	-0.63	0.540170
Comprehensibility & MPhA score	15	-0.16	-0.58	0.574220
Correctness & MPhA score	15	-0.24	-0.89	0.388193

Summing up, the performed tests failed to show any significant correlations between foreign accent, comprehensibility, correctness and, on the other hand, the participants' composite score of metaphonological awareness in three groups (B, C, D), with the exception of Group A, in which the generated coefficients indicated moderate correlations between the investigated pairs of variables.

Finally, a joint analysis for all the groups did not point to any significant correlations between the investigated aspect of perceived pronunciation performance in L3 evaluated by means of accentedness ratings and the participants' composite score for metaphonological awareness ($p>0.05$) (see Table 163).

Table 163. Spearman rank correlation – joint analysis

Pairs of variables	N	R_S	t(n-2)	p
Foreign accent & MPhA score	83	0.19	1.72	0.088557
Comprehensibility & MPhA score	83	0.14	1.24	0.217653
Correctness & MPhA score	83	0.18	1.66	0.100102

6.4.3. Correlation between metaphonological awareness and participants' variables

Further Spearman rank correlation tests were performed to investigate the existence of potential correlations between metaphonological awareness measured by means of the composite score (MPhA) and the participants' variables including age, years of formal training in L2 and L3, age of onset of learning of L2 and L3, proficiency level in L2 and L3, self-evaluation of overall L3 competence and L3 pronunciation, the number of foreign languages known, a stay in an L3-speaking country, and the year of studies.

As far as Group A with L3 French is concerned, significant moderate correlations ($p<0.01$) were found for the following variables: the years of formal training in L3 French (L3_YFT) and MPhA ($R_S=0.56$); the onset age of learning L3 French (L3_AOL) and MPhA ($R_S=-0.57$); proficiency level in L2 English (L2 Prof) and MPhA ($R_S=0.54$); proficiency level in L3 French (L3 Prof) and MPhA ($R_S=0.48$); and the year of studies and MPhA ($R_S=0.51$), (see Table 164).

The generated correlations demonstrated that the more years of formal training in the L3 the participants had, the higher their level of proficiency in both L2 and L3 was, and the higher their year of studies, the more phonologically aware they were. Moreover, there was an inverse relationship with the onset age of learning indicating that the earlier the participants started learning the L3, the higher their metaphonological awareness proved to be.

Table 164. Spearman rank correlations for Group A with L3 French

Pairs of variables	N	R_S	t(n-2)	p
L2_YFT & Composite score	27	-0.02	-0.09	0.928978
L2_AOL & Composite score	27	0.15	0.78	0.443596
L3_YFT & Composite score	27	**0.56**	3.40	**0.002261**
L3_AOL & Composite score	27	**-0.57**	-3.48	**0.001863**
N_TOTAL_Ln & Composite score	27	0.13	0.66	0.516487
Stay in L3 & Composite score	10	0.37	1.14	0.286314
Age & Composite score	27	0.27	1.39	0.175991
Eval L3 compet & Composite score	27	0.38	2.04	0.051899
Eval L3 pron & Composite score	27	0.34	1.79	0.085543
L2_Prof & Composite score	27	**0.54**	3.23	**0.003483**
L3_Prof & Composite score	27	**0.48**	2.72	**0.011745**
Year & Composite score	27	**0.51**	2.97	**0.006437**

In the case of Group B, significant moderate to high correlations ($p<0.05$) were found for the following variables: age and MPhA (R_S= 0.42); the level of proficiency in L2 English (L2 Prof) and MPhA (R_S=0.43); year of studies and MPhA (R_S=0.63), (see Table 165). As it transpires from the conducted correlations, the older the participants, the higher the MPhA composite score; the more advanced they were in L2 English, the higher the MPhA composite score; the higher the year of studies, the more phonologically aware the participants proved to be on the basis of the MPhA composite score.

Table 165. Spearman rank correlations for Group B with L3 German

Pairs of variables	N	R_S	t(n-2)	p
L2_YFT & Composite score	25	0.03	0.15	0.879720
L2_AOL & Composite score	25	0.02	0.09	0.926796
L3_YFT & Composite score	25	0.17	0.83	0.414901
L3_AOL & Composite score	25	0.00	0.01	0.994876
N_TOTAL_Ln & Composite score	25	0.08	0.38	0.704516
Stay in L3 & Composite score	11	-0.23	-0.72	0.487120
Age & Composite score	25	**0.42**	2.25	**0.034397**
Eval L3 compet & Composite score	25	-0.04	-0.19	0.851646

Eval L3 pron & Composite score	25	0.00	0.01	0.990681
L2_Prof & Composite score	25	**0.43**	2.26	**0.033702**
L3_Prof & Composite score	25	0.04	0.20	0.845232
Year & Composite score	25	**0.63**	3.92	**0.000691**

Nearly identical results were generated in Group C with L3 English/L2 German, in which significant moderate correlations ($p<0.05$) were found for the same variables as in Group B, i.e. age and MPhA (R_S= 0.47); the level of proficiency in L2 German (L2 Prof) and MPhA (R_S=0.46); the year of studies and MPhA (R_S=0.55), (see Table 166). All in all, the results demonstrated that the level of advancement in L2 German, the participants' age and advancement in terms of the years of their studies correlated positively with a higher level of metaphonological awareness.

Table 166. Spearman rank correlations for Group C with L3 English/L2 German

Pairs of variables	N	R_S	t(n-2)	p
L2_YFT & Composite score	29	0.18	0.93	0.359724
L2_AOL & Composite score	29	0.26	1.40	0.171596
L3_YFT & Composite score	29	-0.04	-0.21	0.839020
L3_AOL & Composite score	29	0.28	1.51	0.141584
N_TOTAL_Ln & Composite score	29	0.20	1.06	0.298895
Stay in L3 & Composite score	7	0.41	1.01	0.358817
Age & Composite score	29	**0.47**	2.76	**0.010221**
Eval L3 compet & Composite score	29	0.05	0.28	0.780321
Eval L3 pron & Composite score	29	-0.13	-0.70	0.487990
L2_Prof & Composite score	29	**0.46**	2.72	**0.011353**
L3_Prof & Composite score	29	-0.01	-0.05	0.956813
Year & Composite score	29	**0.55**	3.44	**0.001884**

In the case of Group D with L3 English/L2 French, the only statistically significant correlation was observed between the metaphonological awareness (MPhA) composite score and one of the participants' variables. The number of foreign languages the participants' knew correlated moderately with their level of metaphonological awareness (R_S=0.41, $p<0.05$), see Table 167.

Table 167. Spearman rank correlations for Group D with L3 English/L2 French

Pairs of variables	N	R_S	t(n-2)	p
L2_YFT & Composite score	26	-0.31	-1.58	0.126724
L2_AOL & Composite score	26	0.28	1.43	0.166941
L3_YFT & Composite score	26	-0.03	-0.16	0.872060
L3_AOL & Composite score	26	-0.12	-0.59	0.559842
N_TOTAL_Ln & Composite score	26	**0.41**	2.17	**0.040038**
Stay in L3 & Composite score	9	0.03	0.07	0.945843
Age & Composite score	26	-0.14	-0.68	0.502002
Eval L3 compet & Composite score	26	-0.18	-0.90	0.378436
Eval L3 pron & Composite score	26	0.06	0.30	0.766681
L2_Prof & Composite score	26	0.03	0.17	0.869391
L3_Prof & Composite score	26	-0.03	-0.16	0.870767
Year & Composite score	26	0.16	0.78	0.442909

Finally, a joint analysis was performed to investigate relations between the metaphonological awareness (MPhA) composite scores and the participants' variables for all the groups treated jointly. Significant weak to moderate correlations ($p<0.05$) were found for the following variables: age and MPhA composite score ($R_S=0.24$); the total number of foreign languages known (N_TOTAL_Ln) ($R_S=0.24$); the level of proficiency in L2 (L2 Prof) and the MPhA composite score ($R_S=0.38$); the year of studies and the MPhA composite score ($R_S=0.46$), see Table 168. Recapitulating, the joint correlations demonstrated, on the whole, that the participants were found to be more metaphonologically aware when they were older, more proficient in their respective L2s, more advanced in terms of the years of their studies and when they knew more foreign languages.

Table 168. Spearman rank correlations – joint analysis for all groups

Pairs of variables	N	R_S	t(n-2)	p
L2_YFT & Composite score	107	-0.03	-0.30	0.763967
L2_AOL & Composite score	107	0.09	0.92	0.359782
L3_YFT & Composite score	107	0.08	0.79	0.430565
L3_AOL & Composite score	107	-0.07	-0.75	0.457309

N_TOTAL_Ln & Composite score	107	**0.24**	2.48	**0.014828**
Stay in L3 & Composite score	107	0.02	0.22	0.827077
Age & Composite score	107	**0.24**	2.52	**0.013116**
Eval L3 compet & Composite score	107	0.05	0.54	0.588181
Eval L3 pron & Composite score	107	-0.01	-0.12	0.902183
L2_Prof & Composite score	107	**0.38**	4.18	**0.000060**
L3_Prof & Composite score	107	0.07	0.77	0.445336
Year & Composite score	107	**0.46**	5.37	**0.000000**

6.5. Discussion

RQ 1: What forms of metaphonological awareness are manifested by L3 learners? To what extent are the participants able to modify their phonetic output in L3?

The first research question concerned the evidence of metaphonological awareness demonstrated by the participants. The conducted analysis of qualitative and quantitative data collected through the retrospective and introspective verbal protocols showed that various forms of MPhA were manifested at the level of noticing as well as at the level of understanding and metacognition, reflecting the present concept operationalisation and Schmidt's (1990, 1993) categorization of awareness. The manifestations of MPhA at the level of noticing included drawing attention to phonetic features of performance in L3, intentional focus on auditory forms and articulatory gestures, noticing specific problems in L3 pronunciation as well as modifying mispronunciations through ad hoc self- repairs and post hoc self- corrections. Other forms of MPhA manifested through understanding and metacognition consisted of conscious analysis of L3 oral performance, different attempts to formulate phonological rules, explanations of specific language behaviour as well as a considerable degree of self-awareness and self-reflection reflected in the metacognitive comments.

The observed types of metaphonological awareness could also be interpreted in terms of Gombert's (1992) distinction between epilinguistic and metalinguistic awareness. The former category featured spontaneous ad hoc self- repairs during the actual reading performance in the L3, yet the instances thereof were rather few. The latter category exhibited examples of post-hoc self- corrections of the oral performance, and reflective analysis of L3 pronunciation on the part of the participants. It appears that the instances of metalinguistic awareness were more frequent, thus pointing to a more conscious nature of the participants' metaphonological awareness.

The second part of this research question aimed to explore to what extent the participants were able to modify their phonetic output in L3. To this end, instances of self-correction of L3 pronunciation performance were analysed with respect to their success rate. As shown in the joint analysis of the results, on average 50% of the phonetic modifications provided by the participants were correct and more advanced learners tended to be more successful in their repairs and performed fewer incorrect modifications, as expected.

RQ 2: How complex is the conducted metaphonological analysis of L3 performance?

The following research question concerned the complexity of the performed metaphonological analysis. It was assessed by adapting Roehr's (2006) metalinguistic awareness categorization to the present study in order to distinguish three levels of complexity, ranging from evidence of a low-level noticing by e.g. attentional focus on the auditory form; through a mid level characterized by metalinguistic description or explanation; to a high level of explicit metalinguistic knowledge evidenced in the use of appropriate metalanguage.

The results demonstrated that the majority of the introspective and retrospective verbal protocols was classified as belonging to the low complexity level, thus pointing to the prevalence of more implicit forms of awareness, reflecting mostly the 'noticing' dimension of Schmidt's (1990, 1993) hypothesis. Furthermore, a considerable percentage of the responses corresponded to the medium complexity level, evidencing a more explicit awareness, manifested through the participants' attempts at analyzing and accounting for their phonetic performance in L3. This, in turn, related to the 'understanding' dimension of Schmidt's proposal. However, the high complexity level of awareness, which involved explicit verbalizations of metalanguage, was quite marginal.

Trying to account for the findings with regard to Bialystok's (2001) model of developing analysed knowledge and control, we saw evidence of largely intuitive knowledge manifested through numerous low complexity metacomments, complemented by elements of conscious knowledge (cf. instances of mid complexity level) and rather limited evidence of fully analysed knowledge in the third language (i.e. a limited number of high complexity comments). With respect to the control component that is held responsible for the automatisation of linguistic performance in an L3, the occurrence of relatively infrequent instances of ad hoc self-repairs as opposed to much more frequent post hoc self- corrections may point to a rather low degree of automatic control of oral performance in L3. On the other hand, a number of the metacognitive comments provided in the introspective protocols indicated that the participants were highly aware of the problems of control evidenced e.g. through articulatory control, code-switching and the perception of foreign accentedness in the third language.

The findings only partially support Mora et al.'s (2014) stipulations about the implicit nature of phonological awareness in a foreign language, which they consider essential for phonological language acquisition. In the present study, also elements of explicit metaphonological knowledge are attested, although less frequently. They are manifested through the verbalisations of phonological rules, reported CLI and metacomments on phonological differences between L1, L2 and L3 systems, thus evidencing a deeper knowledge of the target language structure.

RQ 3: Is there evidence of cross-linguistic awareness in L3 learners? What is the source language for CLI in L3 acquisition, according to the participants?

The third research question aimed to investigate evidence of cross-linguistic awareness. It relied on Jessner's (2006) stipulation that multilingual subjects demonstrate an additional component of metalinguistic awareness related to the interactions within their multilingual repertoires. The present findings provide further evidence for the cross-linguistic awareness of the multilingual participants, reflected in their introspections about perceived influences and interactions between various language systems. The declared sources of cross-linguistic influence featured both native and non-native language systems, thus attesting L1-to-L3 and L2-to-L3 transfer.

The majority of the participants reported that their second language (L2) prevails as the source of transfer in the third language. This corroborated, to a large extent, the assumption of the foreign language effect', or 'L2 status', commonly acknowledged in the recent L3 literature. In accordance with its tenets, the second language constitutes a prevailing source of cross-linguistic influence in the acquisition of a subsequent language and there is a tendency to activate the L2 in L3 production (cf. Cenoz 2001; De Angelis 2007, Hammarberg and Hammarberg 2005). Since a smaller percentage indicated their L1 as the main source of cross-linguistic phonological influence in the acquisition of their respective third language, it provided weak support for the traditionally held belief that the articulatory motor routines established in the mother tongue determine the acquisition of any subsequent language systems, be it L2 or L3 (Ringbom 1987). The third category of responses indicating combined cross-linguistic influence provided considerable support to De Angelis' (2007) proposal of a combined cross-linguistic influence that was, in turn, corroborated by the results of previous studies on third language phonology (cf. Benrabah 1991, Blank and Zimmer 2009, Hammarberg and Hammarberg 2005, Wrembel 2010, 2012, Wunder 2010).

In the qualitative analysis of the reported instances of CLI, several conditioning factors were identified by the participants, including the route of acquisition, frequency or recency of the use of particular languages, or typological closeness. Particularly noteworthy is that the majority of the factors acknowledged by the participants as conditioning the process of cross-linguistic influence, based on their intuitions and introspections, reflected the exact set of determiners that were identified in the third language acquisition literature (cf. Cenoz, 2001; De Angelis, 2007; Hammarberg, 2009). Interestingly, the majority of the metacognitive comments on multilingual advantage in subsequent language learning corresponded closely to the arguments put forward in the third language acquisition literature (cf**.** Cenoz 2003, Cenoz and Jessner, 2000, De Angelis 2007, Gut 2010, Tremblay 2010). Nonetheless, further research is still necessary to confirm the facilitative effect of bilingualism on L3 acquisition, and particularly to investigate if the multilingual advantage is specific for phonology rather than for general language proficiency.

RQ 4: Is there a correlation between the MPhA composite score and accent rating parameters in L3?

The following research question addressed the issue whether the meta-phonological awareness composite score was related to the perceived pronunciation performance in L3. The conducted statistical analyses did not demonstrate significant correlations between accentedness, compre-hensibility, correctness and the participants' composite score of meta-phonological awareness in three groups (B, C, D), with the exception of Group A, in which the generated coefficients indicated moderate correla-tions between the pairs of variables under investigation. The findings suggest that metaphonological awareness as measured by the MPhA com-posite score does not necessarily contribute to better comprehensibility, foreign accentedness or pronunciation correctness based on the perceptual ratings, as the evidence for the existence of such a relation was limited to one group only.

Previous research by Venkatagiri and Levis (2007) indicated the existence of a positive correlation between phonological awareness measures and rated speech comprehensibility. Similarly, Kennedy and Trofimovich (2010) reported a positive correlation between the L2 pronunciation ratings of accentedness, comprehensibility and fluency and the number of qualitative (but not quantitative) language awareness comments, i.e., higher pronunciation ratings were associated with a greater number of qualitative language awareness comments related to using pronunciation to convey the intended message. The results of the two studies were only partially confirmed by the present findings as a similar correlation was found only in one of the investigated groups (A), whereas the remaining groups failed to determine such an interrelation between the MPhA score and perceived pronunciation performance in L3 evaluated by means of the ratings of foreign accentedness, comprehensibility and correctness. However, as far as Mora et al's (2014) results are concerned, their measure of L2 phonological awareness expressed through mimicking L2-accented speech in L1 Spanish did not show any significant correlations with the performed foreign accentedness ratings, similarly as in the current studies.

RQ 5: Is there an interdependence between the MPhA composite score and the participants' proficiency level as well as other variables?

The final research question aimed to explore potential correlations between metaphonological awareness and the participants' variables, including their proficiency level. It was expected that a higher proficiency in the third language would result in a more enhanced metaphonological awareness reflected in the L3 oral performance. Separate analyses were performed to investigate the impact of the L3 proficiency level on the participants' performance on different tasks. A detailed analysis of the data confirmed the initial assumption since, for instance, more advanced proficiency subgroups performed more correct modifications of their L3 speech (in Groups A, B, C) and were able in general to notice more problems with L3 pronunciation. Moreover, higher rates of reported CLI were observed in the more proficient subgroups for two groups (B, C). In terms of the complexity level of the generated comments and metacomments, the proficiency group effect was reported in three out of four groups (i.e. A, B, C) with the more advanced subgroups outperforming the less proficient ones on at least two complexity levels. As far as the formulation of phonological rules is concerned, an L3 proficiency group effect was evidenced in three out of four groups (i.e. A, B, C) in that more proficient learners were able to provide more instances of phonological rules. Recapitulating, the assumption of an L3 proficiency group effect on the degree of metaphonological awareness was substantiated in the great majority of the parameters when the performance of the proficiency subgroups was compared. Interestingly, the proficiency group effects were not reported in the author's previous exploratory study (Wrembel 2015) as in this case the participants' levels of proficiency in L3 Polish were probably not divergent enough to demonstrate any significant differences.

The second part of the research question investigated the relationship between metaphonological awareness and other participants' variables. The performed Spearman rank correlations demonstrated that several factors were correlated with the MPhA composite score, including the participants' age, the level of proficiency in the second and third language, the year of studies, and the number of foreign languages known by the participants. Previous studies rarely investigated any effects of learners' variables on the level of phonological awareness. For instance, Venkatagiri and Levis (2007) reported a positive correlation between phonological awareness and phonological short term memory. On the other hand, Kennedy and Trofimovich (2010) found that the participants who had a high degree of phonological awareness reflected in the number of qualitative language awareness comments were found to report the most extensive out-of-class listening to the L2 input. Mora et al. (2014) did not account for any participants' variables in their study at all.

This part of the project, i.e. Study 3, provides a particularly welcome contribution to a better understanding of the field, as research into phonological awareness has been rather scarce to date (see Section 3.2.9 for an overview). The present findings have largely confirmed the results generated in the author's preliminary studies into metaphonological awareness conducted on smaller pools of subjects with L3 German and L3 French (Wrembel 2013) as well as an innovative investigation of L3 Polish (Wrembel 2015). The previous studies allowed to develop and test a complex coding system proposed by the author to analyse quantitatively and qualitatively datasets generated through verbal protocols. Further, they served as a basis for designing a formula for quantifying a composite measure of metaphonological awareness that was put forward for the first time in the present large scale investigation.

Previous related research focused primarily on tapping implicit metalinguistic knowledge related to L3 phonological acquisition (e.g. Venkatagiri and Levis 2007, Mora et al. 2014), whereas the present study aimed to expand this perspective to include the explicit component as well. Since, as stipulated by White and Ranta (2002), metalinguistic performance involves both implicit and explicit knowledge, in the current investigation introspective and retrospective oral protocols were used as a data collection procedure to reflect these two aspects of metaphonological awareness. Since the participants of this series of studies reported having a general background in linguistics and formal training in L2 pronunciation, but no or little formal instruction in L3 pronunciation, it was interesting to explore how they apply this general knowledge to the acquisition of a new phonological system.

Unlike Venkatagiri and Levis (2007), the present investigation did not rely on measures of phonological awareness developed for first language acquisition, including phoneme manipulation, blending, deletion, segmentation and sequencing as well as rhyming and alliteration abilities related to literacy development. The application of such measures was questioned on methodological grounds by Mora et al. (2014), who maintain that the foreign language acquisition perspective requires different means, as defining phonological awareness (PhonA) as “the ability to segment and manipulate sounds seems more adequate and psychologically real in the domain of L1 literacy acquisition than in second language (L2) acquisition” (Mora et al. 2014: 58).

Some similarities can be observed between the present study and that of Kennedy and Trofimovich (2010), which examined the relationship between the quality of L2 pronunciation as measured by ratings of accentedness, comprehensibility and fluency, on the one hand, and language awareness, on the other. Kennedy and Trofimovich (2010) assessed L2 pronunciation awareness based on dialogue journal entries and distinguished between quantitative or qualitative comments, which may be somehow related to the verbal protocols data collection performed in the present study. However, the conducted analysis of data differed significantly, as the above mentioned authors relied on Benson and Lor’s (1999) analytical framework of the dual conceptions of learning, and limited their observations mostly to how pronunciation awareness assists language acquisition (i.e. quantitative awareness) or how language works to convey meaning (i.e. qualitative awareness).

Although language awareness is usually defined as explicit knowledge about language (cf. the Association of Language Awareness 2012), Mora et al. (2014) and Kivistö-de Souza (2012) proposed a different understanding of phonological awareness in the second language that involves implicit knowledge about the phonological system of the target language manifested at the segmental, suprasegmental and phonotactic levels. According to this understanding, L2 phonological awareness develops mainly implicitly based on learners’ capability of discerning differences between sounds in their L1 and L2, extracting L2-specific phonetic and phonological regularities, which, consequently, leads to modifications of L2 phonological representations (cf. Mora et al. 2014: 58). In an attempt to apply Schmidt’s noticing hypothesis to their approach, the authors claim that in L2 phonology “awareness at the level of *noticing* is a requirement for the acquisition of a target phonological structure, whereas awareness at the level of *understanding* can be beneficial but is not essential to L2 phonological acquisition” (Mora et al. 2014: 58). Embracing another perspective, Venkatagiri and Levis (2007:

265) define phonological awareness as "conscious knowledge of the sounds, syllable structure, phonotactics and prosody of the target language". The stance of the present author is, however, that both the implicit and explicit knowledge constitute important elements of the metaphonological awareness, as evidenced in the present findings generated from the introspective and retrospective verbal protocols.

The findings provide support for Schmidt's (2009) noticing hypothesis, as the participants with a higher metaphonological awareness exhibited also a higher proficiency level in their L2 and, to some extent, in the L3, thus testifying that conscious awareness is a conditioning factor for learning. Schmidt's noticing hypothesis stipulates that language awareness can be represented as a continuum of different levels of consciousness, encompassing perceiving as the lowest end of the continuum, through noticing to understanding, as the most complex manifestation of awareness. The present evidence is reflected in the participants' perceiving mismatches between L3 output and input, noticing discrepancies in the auditory form between their L3 production and target forms, attending to the phonetic properties of speech in the L3, analyzing phonological structures as well as showing an understanding of the rules and processes behind them through explicit verbalizations.

As far as potential pedagogical implications are concerned, the findings relate closely to the focus on form (FonF) framework of instruction, according to which learner's attention should be mostly allocated to language form by means of awareness raising and developing conscious noticing (cf. Long and Robinson 1998 for a detailed discussion). The superiority of such an approach over a traditional focus on forms syllabus was advocated by Doughty and Williams (1998), who further proposed that the FonF framework could be applied to different types of language forms, including phonology.

The results provide also evidence for the claims promoted by multilingual educators and scholars investigating third language acquisition, for instance Jessner (2006), in accordance with which prior language knowledge is activated in the foreign language classroom in the form of multiple comparisons between language systems and an acknowledgement of cross-linguistic influence. A pedagogical proposals tailored specifically towards third language pronunciation was first put forward by Marx and Mehlhorn (2010). They suggested developing awareness of the impact of factors other than L1 transfer as well as taking advantage of the potential for positive phonetic interference from other non-native languages. Further techniques geared at enhancing metaphonological awareness in third and additional language learning proposed by the present author involve, among others, making multilingual learners aware of their broadened phonetic repertoires and language learning strategies; directing learners' attention to the mismatch between input and output; developing active listening to phonetic input in a foreign language; encouraging conscious self-analysis of one's oral performance; or reinforcing cross-linguistic awareness to facilitate the acquisition of further languages (cf. Wrembel and Sypiańska 2014 for a more detailed discussion). Future research should, however, address the question of the effectiveness of pedagogical interventions directed at raising phonological awareness in a foreign language.

6.6. Conclusion

The present investigation was intended to shed more light on the unexplored area of awareness and noticing in the acquisition of foreign language phonology. It aimed to explore how multilingual learners attend to and modify their phonetic output in the third language, consciously analyse it and comment on their oral performance in L3. The study adopted a multilingual perspective to reflect the complexity and importance of metaphonological awareness in the context of multiple language acquisition.

The participants of this series of studies succeeded at demonstrating various forms of implicit and explicit metaphonological awareness. They were able to notice deviations from target-like pronunciation in their L3 performance (e.g. instances of self-correction) and perceive degrees of correctness in their speech performance (e.g. noticing problems with L3 pronunciation). Moreover, they demonstrated sensitivity to phonological differences between the language systems at their disposal, both at the segmental and suprasegmental levels, manifested through the performed conscious phonetic analysis or cross-linguistic auditory or articulatory comparisons.

The results indicate, in accordance with Wrembel's (2013, 2015) preliminary investigations, that the L3 learners can achieve a certain level of metaphonological awareness in spite of the lack of, or only limited, formal training in the phonology of the third language. Undeniably, the participants of the study relied to some extent on their general linguistic training as well as pronunciation instruction provided in the course of their L2 acquisition. It would be interesting to expand the study to include linguistically naïve learners and to tap their implicit phonological awareness; however, this is beyond the scope of the present study.

Taking into account the fact that the majority of theoretical models, including the most influential ones by Bialystok (1994, 2001), Schmidt (1990, 1993, 2001) or Dziubalska-Kołaczyk (1990) advocate consciousness as a necessary condition for learning a foreign language, the exploration of metaphonological awareness appears to be of paramount importance. The present contribution is innovative not only because it addresses an unexplored domain of cognitive factors of awareness and noticing involved in phonological acquisition but also since it does so from a third language acquisition perspective, further evidencing metalinguistic awareness as a key component of multilingual competence.

Chapter 7

GLOBAL ANALYSIS

7. Global analysis of cross-linguistic influence in L3

In this chapter a global analysis of the results generated in the three series of studies, as discussed in Chapters 4-6, will be presented. The chapter will consist of four parts including: across group comparisons, across studies comparisons, the verification of the research hypotheses and final conclusions.

7.1. Across group comparison

In order to address further the research hypotheses stipulated in Chapter 2 (see 2.5.1), several across-group comparisons of the results of the three performed series of studies were carried out. The conducted analysis involved the following four across-group comparisons corresponding closely to the research hypotheses:

– A vs. B, i.e. shared L2 English, different L3s: French and German
– C vs. D, i.e. shared L3 English, different L2s: German and French
– B vs. C, i.e. a typologically related mirror pair: L3 German/L2 English vs. L3 English/L2 German
– A vs. D, i.e. a typologically unrelated mirror pair: L3 French/L2 English vs. L3 English/L2 French.

The across-group comparisons were performed for the following parameters investigated in the respective three studies:

– accentedness ratings composite score (FA) – Study I
– L1 identification (as L1, as L2, as other languages) – Study I
– goodness of fit of L3 to control VOT values – Study II
– VOT values for /p, t, k/ for particular languages (English, French, German and Polish) – Study II
– metaphonological awareness composite score (MPhA) – Study III.

The comparisons for the pairs of groups for the variable parameters were performed by means of three different statistical tests, depending on which conditions were met. Firstly, the t-Student Test for independent samples was run when the distribution of variables under examination was normal and the condition of variance homogeneity was met. The results of the test were described by means of the following parameters: (1) t – t test value, (2) df – degree of freedom, (3) p – probability level. Secondly, the Cochran-Cox Test, i.e. a t test for nonhomogeneous variances for two unrelated samples was performed when the variables under examination were normally distributed; however, the condition of variance homogeneity was not met. The results of the test were described by means of the following parameters: (1) t_{C-C} – t test value for nonhomogeneous variances, (2) df – degree of freedom, (3) p – probability level. Thirdly, the Mann-Whitney Test, i.e. a nonparametric test for two unrelated samples was calculated when one of the groups did not follow a normal distribution. The results of the test were described by means of the following parameters: (1) U – Mann-Whitney Test value (used for small samples below 20), (2) Z corr – Mann-Whitney Test corrected value used for tied ranks and samples above 20, (3) p – probability level.

7.1.1. Comparisons for accentedness ratings

A composite score of accentedness ratings was calculated based on the three rating parameters, including foreign accentedness, comprehensibility and pronunciation accuracy. The descriptive statistics for all the groups is presented in Table 169.

Table 169. Descriptive statistics for accentedness ratings across groups

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	28	3.98	1.21	4.33
B	L3 Ger_L2 Eng	25	3.98	0.78	4.02
C	L3 Eng_L2 Ger	17	3.56	0.60	3.68
D	L3 Eng_L2 Fre	16	4.13	0.53	4.17

The highest accentedness ratings were reported for group D with L3 English_L2 French (M=4.13), and the lowest for group C with L3 English_L2 German (M=3.56). The remaining two groups had very similar scores (for group A with L3 French_L2 English M=3.98, group B with L3 German_L2 English M=3.98).

The performed across-group comparisons showed one statistically significant difference between the investigated pairs of groups, i.e. for the C vs. D groups with shared L3 English and different L2s. The mean value of the accentedness rating composite score was lower in group C with L3 English_L2 German than in group D with L3 English_L2 French, and the difference is statistically significant according to the conducted t-test ($t=-2.91$, $df=31$, $p=0.006$). All the remaining across group comparisons did not point to any significant differences as far as the accentedness score is concerned (A vs. B: $Zcc=0.53$, $p>0.05$; B vs. C: $t=1.91$, $df=40$, $p>0.05$; A vs. D: $U=217$, $p>0.05$).

7.1.2. L1 identification

The second part of accentedness ratings in Study I consisted of an L1 identification based on the samples of L3 speech. The possible options included correct L1 identification as L1 Polish, L1 identification as L2, and L1 identification as other languages. The sections below discuss the results for particular categories with respect to the comparative group analysis.

The descriptive statistics for L1 identifications as L1 Polish are presented in Table 170. As can be seen, the highest correct identification rate was reported in the case of group B with L3 German_L2 English (M=41.6), and the lowest in group C with L3 English_L2 German (M=29.2), with groups A and D scoring at a comparable level (M=36.2 and 37.8 respectively).

Table 170. Descriptive statistics of correct L1 identification as L1

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	28	36.2	12.6	34.1
B	L3 Ger_L2 Eng	25	41.6	12.7	45.0
C	L3 Eng_L2 Ger	17	29.2	11.6	26.1
D	L3 Eng_L2 Fre	16	37.8	13.2	39.1

The performed across group comparisons demonstrated only one statistically significant difference between the group pairs, i.e. B vs. C – a typologically related mirror pair. The mean value of the L1 identification as L1 Polish was higher in the L3 German_L2 English group than in the L3 English_L2 German group, and the results of the conducted t-test were statistically significant ($t=3.22$, $df=40$, $p=0.002$). There were no significant differences between the remaining group comparisons, i.e. A vs. B ($t=-1.55$, $df=51$, $p>0.05$), C vs. D ($t=-2.0$, $df=31$, $p>0.5$) and A vs. D ($t=0.39$, $df=42$, $p>0.05$).

As far as the identification of the native tongue of the L3 speakers as the respective L2 is concerned, the results for all the groups are presented in Table 171.

Table 171. Descriptive statistics of L1 identification as L2

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	28	8.1	11.4	4.5
B	L3 Ger_L2 Eng	25	10.6	10.2	10.0
C	L3 Eng_L2 Ger	17	17.6	21.3	8.7
D	L3 Eng_L2 Fre	16	14.1	18.2	10.9

A comparative analysis demonstrated that there were no statistically significant differences between the groups. The median was the lowest for group A with L3 French_L2 English ($\tilde{x}=4.5$), intermediate for group C with L3 English_L2 German ($\tilde{x}=8.7$), and the highest for group B with L3 German_L2 English ($\tilde{x}=10$) and group D with L3 English_L2 French ($\tilde{x}=10.9$).

Finally, the identification of the native tongue of the L3 speakers as languages other than their respective L1 or L2 yielded results in the range of 48-56%. A higher percentage of identifications was reported for groups A and C, with a slightly lower level in the case of groups B and D (see Table 172).

Table 172. Descriptive statistics of L1 identification as other languages

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	28	55.7	13.3	59.1
B	L3 Ger_L2 Eng	25	47.8	9.6	45.0
C	L3 Eng_L2 Ger	17	53.2	16.2	60.9
D	L3 Eng_L2 Fre	16	48.1	16.5	47.8

The results of the performed t-tests showed a significant difference only between group A with L3 French_L2 English, and group B with L3 German_L2 English ($t=2.45$, $df=51$, $p<0.05$) with the former mean values for identification as "other" being significantly higher ($M=55.7$) than the latter ($M=47.8$). No significant differences were found for the remaining groups as far as L1 identification as other languages is concerned.

7.1.3. Comparison of the L3 VOT goodness of fit to control

As far as the comparison of the goodness of fit of the L3 VOT values to the control values is concerned, all the groups departed considerably from the target VOT values. The descriptive statistics for all the groups is presented in Table 173. In the case of group A with L3 French_L2 English, the VOT durations were on average 25% higher than the target French VOT, whereas for the remaining groups the L3 VOT fell below the target values (for B on average 22% below the target German VOT, for C and D 32% - 33% below the target English VOT).

Table 173. Descriptive statistics for L3 VOT goodness of fit to control

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	2.54	3.27	1.72
B	L3 Ger_L2 Eng	26	-2.21	1.90	-1.97
C	L3 Eng_L2 Ger	33	-3.21	1.49	-3.44
D	L3 Eng_L2 Fre	28	-3.30	2.15	-3.56

The comparative across-group analysis demonstrated significantly different results in the case of three pairs of groups, i.e. A vs. B, B vs. C, and A vs. D. In the A vs. B group comparison, the mean value of the L3 VOT goodness of fit to the L3 control VOT in group L3 French_L2 English ($M=2.54$) was higher than in the group L3 German_L2 English ($M=-2.21$) and this difference was statistically significant ($Zcorr=5.64$, $p<0.05$).

For the B vs. C group comparison, the mean value of the L3 VOT goodness of fit to the L3 control VOT was higher in the L3 German_L2 English group ($M=-2.21$) than in the L3 English_L2 German group ($M=-3.21$) and the difference was statistically significant ($t=2.26$, $df=57$, $p=0.027$). In the comparison between the A vs. D groups, the mean value of the L3 VOT goodness of fit to the control VOT is lower in the L3 English_L2 French group (M=-3.3) than in the L3 French_L2 English group (M=2.54) and this difference is statistically significant (Zcorr=-6.27, p<0.05). There was no statistically significant difference in the L3 VOT goodness of fit to control for the C vs. D group comparison, i.e. with shared L3 English and different L2s (Zcorr=0.61, p>0.05). Interestingly, the goodness of fit ratio for L3 English was in both cases comparable in spite of the different L2, i.e. French and German, which differ also in terms of their respective VOT durations.

A separate category of the analysis was the comparison of the goodness of fit to the control VOT performed for the respective L2 values. Table 174 presents the descriptive statistics for all the groups for this comparison. As can be seen, three out of the four groups (A, B, D) demonstrate a goodness of fit ratio for L2 VOT values that are very close to the target durations (from 2% to 5% deviations). Only in the case of group C with L3 English_L2 German, do the VOT values for L2 German deviate from the control VOT on average by 29%.

Table 174. Descriptive statistics for L2 VOT goodness of fit to control

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	0,59	2,30	0,43
B	L3 Ger_L2 Eng	26	0,21	2,79	0,11
C	L3 Eng_L2 Ger	33	-2,89	1,37	-3,35
D	L3 Eng_L2 Fre	28	-0,45	1,96	-0,61

As far as the performed across group comparisons are concerned, statistically significant differences were demonstrated for the C vs. D and B vs. C comparisons. In the former case, the mean for the L2 VOT_FIT control was lower in the L3 English_L2 German group than in the L3 English_L2 French group (Zcorr=-4.78, p<0.05). In the latter case, the mean for the L2 VOT_FIT control was higher in the L3 German_L2 English group than in the L3 English_L2 German group (Zcorr=4.84, p<0.05). The comparisons for A vs. B and A vs. D did not show any statistically different L2 VOT goodness of fit ratio. Summing up, the values for L2 English and L2 French were close to the target control VOT durations, the only exception being the L2 German VOT – which departed from the control values.

7.1.4. Comparison for MPhA composite score

The descriptive statistics for the comparative analysis of the metaphonological awareness composite score (MPhA) is presented in Table 175.

Table 175. Descriptive statistics for MPhA across groups

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	27	28.3	11.9	29.0
B	L3 Ger_L2 Eng	25	39.5	17.6	40.5
C	L3 Eng_L2 Ger	29	26.3	10.5	25.5
D	L3 Eng_L2 Fre	26	30.8	10.7	27.0

The highest MPhA score was reported for group B with L3 German_L2 English, whereas the lowest for group C with a mirror language pairing, i.e. L3 English_L2 German. The remaining two groups, A and D, displayed a comparable medium score.

The performed across-group comparisons pointed to statistically significant differences for two pairs of groups, i.e. A vs. B and B vs. C. As far as the comparison between groups A and B with shared L2 English and different L3s is concerned, the L3 German group scored better in the metaphonological awareness measure than the L3 French group. The mean value of MPhA composite score is significantly lower in group A (M=28) than in group B (M=39.5) and the difference calculated by means of the t-test is statistically significant (t=-2.70, df=50, p=0.009). Further, the comparison between B and C groups with typologically related mirror pairs: L3 German/L2 English vs. L3 English/L2 German, also demonstrated a statistically significant difference. The mean value of the MPhA composite score is higher in group B with L3 German_L2 English (M=39.5) than in group C with L3 English_L2 German (M=26), and according to the Cochran-Cox test the difference is statistically significant (t_{CC}=3.27, df=38, p=0.002).

The comparisons of the metaphonological awareness scores for the remaining two group pairs did not yield any significant differences, i.e. for C vs. D with shared L3 English and different L2s (t=-1.55, df=53, p>0.05) and for A vs. D – a typologically unrelated mirror pair (t=0.79, df=51, p>0.05).

7.1.5. VOT values for /p, t, k/ for particular languages

As part of the analysis for Study II, the VOT values for /p, t, k/ for particular languages (i.e. English, French, German and Polish) were compared across the groups. The following sections will present the comparative analysis for the four languages separately.

7.1.5.1. English VOT for /p, t, k/

Table 176. Descriptive statistics for English VOT for /p/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	65.0	21.7	62.8
B	L3 Ger_L2 Eng	26	57.5	23.8	55.7
C	L3 Eng_L2 Ger	33	34.1	12.6	32.8
D	L3 Eng_L2 Fre	28	32.3	16.2	28.8

As shown in Table 176, the mean VOT durations for /p/ were higher for groups A and B (M=65 and M=57.5), in which English has the L2 status, than in groups C and D, in which English is the L3 (M=34 and M= 32 respectively).

When compared across the groups, the VOT values for /p did not differ significantly (p>0.05) for English as the L2 or English as the L3. In the case of L2 English, the difference between groups A with L3 French_L2 English (M=65) and B with L3 German_L2 English (M=57.5) was not found to be statistically significant (Student t-test: t=1.3, df=62, p>0.05). Likewise, the comparison of the L3 English groups, i.e. group C with L3 English_L2 German (M=34) and group D with L3 English_L2 French (M=32) did not yield any significantly different results (Man-Whitney test: Z corr=0.9, p>0.05). It follows that the different L2s (French and German) did not exert any significant influence on the L3 English values.

However, across group comparisons between the VOT durations in English as L2 and as L3 did yield significantly different results. The performed Mann-Whitney test indicated a statistically significant difference in the VOT durations between groups B and C (Z corr=4.3, p<0.05) with VOT values for L2 English (M=57.5) being significantly higher than for L3 English (M=34). The comparative analysis between groups A and D yielded similar results (Z=-5.7, p<0.05) with the VOT for /p/ in L2 English (M=65) being significantly longer than in L3 English (M=32).

Table 177 presents the descriptive statistics for the English VOT for /t/. As can be seen, the highest VOT durations were reported in groups A (M=74) and B (M=74), in which English has the L2 status, as opposed to lower VOT values for groups C (M=41) and D (M=41), where it is the L3.

Table 177. Descriptive statistics for English VOT for /t/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	74.2	22.5	76.5
B	L3 Ger_L2 Eng	26	73.8	24.1	79.2
C	L3 Eng_L2 Ger	33	41.1	12.5	39.3
D	L3 Eng_L2 Fre	28	40.8	17.0	39.6

The observed VOT patterns for English /t/ were identical to those reported for English /p/. The performed comparative analysis showed no statistically significant differences between groups A vs. B ($t=0.06$, $df=62$, $p>0.05$) or C vs. D ($Zcorr=0.28$, $p>0.05$). The comparisons of groups with English as the L2 vs. those with English as the L3 generated significant differences; for B vs. C ($tcc=6.29$, $df=35$, $p<0.05$) and for A vs. D ($Zcorr=-5.7$, $p<0.05$).

In turn, Table 178 shows the comparative results for all the groups for the English VOT for /k/.

Table 178. Descriptive statistics for English VOT for /k/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	90.1	15.6	91.8
B	L3 Ger_L2 Eng	26	89.3	19.5	84.3
C	L3 Eng_L2 Ger	33	65.7	11.9	67.0
D	L3 Eng_L2 Fre	28	65.4	19.3	62.5

Identical patterns were observed in the case of the English VOT for /k/ as for the English VOT for /p/ and /t/. The performed comparative analysis showed no statistically significant differences between groups with English as the L2, i.e. A vs. B ($t=0.20$, $df=62$, $p>0.05$) or with English as the L3, i.e. C vs. D ($tcc=0.08$, $p>0.05$). The comparisons of groups with mixed English status, i.e. L2 vs. L3, demonstrated significant differences; B vs. C ($tcc=5.4$, $df=39$, $p<0.05$) and for A vs. D ($t=-5.7$, $df=64$, $p<0.05$).

Across the groups comparisons were performed also jointly for all the plosives. The mean VOT English values did not differ significantly (p>0.05) for the groups with English as the L2 or the groups with English as the L3. In the case of L2 English, the difference between groups L3 French_L2 English (M=76.4) and L3 German_L2 English (M=73.5) was not found to be statistically significant (Student t-test: t=0.59, p>0.05). Likewise, the comparison of the L3 English groups' L3, i.e. Eng_L2 Ger (M=46.8) and L3 English_L2 French (M=46.2) did not yield any significantly different results (Man-Whitney test: Z corr=0.68, p>0.05). It follows that the different L2s (French and German) did not exert any significant influence on L3 English mean VOT values.

However, the comparisons between VOT durations in English as L2 and as L3 did result in significant differences. The performed Cochran-Cox test indicated a statistically significant difference between the mean VOT durations for L3 Ger_L2 Eng vs. L3 Eng_L2 Ger (tcc=5.8, p<0.05) with the VOT values for L2 English (M=73.5) being significantly higher than for L3 English (M=46.8). The comparative analysis between L3 Eng_L2 Fre vs. L3 Fre_L2 Eng yielded similar results (Man-Whitney test; Zcorr=-5.7, p<0.05) with the mean VOT in L2 English (M=76.4) being significantly longer than in L3 English (M=46.2).

7.1.5.2. French VOT for /p, t, k/

The VOT values for French were compared between Groups A and D, where French had a different status, either as the L2 (D) or as the L3 (A). Comparative analyses were performed separately for the individual plosives /p/, /t/, /k/ as well as for the mean French VOT treated jointly (see Table 179).

In the case of the French VOT for /p/, the median was lower in group A with L3 English_L2 French than in group D with L3 French_L2 English and the difference was statistically significant (Zcorr=-4.48, p<0.05). The median for the French VOT for /t/ was lower in group A than in group D and the difference proved also significant (Zcorr=-2.52, p<0.05). Similarly, the mean VOT for French /k/ was significantly lower in group A than in Group D (t=-5, p<0.05). The comparative analysis performed jointly for all the French voiceless plosives also yielded significant differences across the groups (Zcorr=-4.21, p<0.05).

Table 179. Descriptive statistics for French VOT for /p, t, k/

French VOT	Group	Languages	N	Mean	SD	Median
/p/	A	L3 Fre_L2 Eng	28	25.7	7.9	24.7
	D	L3 Eng_L2 Fre	38	39.5	13.6	38.5
/t/	A	L3 Fre_L2 Eng	28	34.6	8.8	33.0
	D	L3 Eng_L2 Fre	38	42.3	13.6	40.4
/k/	A	L3 Fre_L2 Eng	28	52.3	9.9	53.4
	D	L3 Eng_L2 Fre	38	67.9	14.1	69.3

The reported difference may be related to the varying status of French either as the L2 (Group D) or as the L3 (Group A). In the case of the L2 French the VOT durations are shorter and approximate more the target French short lag values. Interestingly, the L3 French VOT values exceed the target ones, and this finding could be interpreted as some evidence of the influence of longer lag VOT values in L2 English on L3 French, thus pointing to the L2 status effect.

7.1.5.3. German VOT for /p, t, k/

Table 180 presents the descriptive statistics for the German VOT values across the groups. As can be seen, the durations in group B with L3 German are on the whole longer than in group C with L2 German.

Table 180. Descriptive statistics for German VOT for /p, t, k/

German VOT	Group	Languages	N	Mean	SD	Median
/p/	B	L3 Ger_L2 Eng	26	41.9	16.6	41.8
	C	L3 Eng_L2 Ger	33	38.7	11.2	35.8
/t/	B	L3 Ger_L2 Eng	26	48.3	16.8	48.3
	C	L3 Eng_L2 Ger	33	37.9	11.7	33.5
/k/	B	L3 Ger_L2 Eng	26	70.3	13.1	70.5
	C	L3 Eng_L2 Ger	33	68.2	11.5	69.8

The VOT values for German were compared between Groups B and C, where German had a different status, either as the L2 (C) or as the L3 (B). Comparative analyses were performed separately for the individual plosives /p/, /t/, /k/ as well as for the mean German VOT treated jointly.

The analysis of the German VOT in Groups B (L3 German_L2 English) and C (L3 English_L2 German) demonstrated a lack of any significant differences for /p/, /k/ and the joint analysis for all the plosives, although in all these cases the VOT values for the L3 German were slightly higher than for L2 German. The only statistically significant differences between groups B and C were found for the German /t/, for which the VOT duration in L3 German was significantly higher than in the L2 German ($Z_{corr}=2.5$, $p<0.05$).

7.1.5.4. Polish VOT for /p, t, k/

Polish had the status of the L1 in all the groups under investigation. Comparative VOT analyses were performed separately for individual plosives /p/, /t/, /k/ as well as for the mean Polish VOT treated jointly to see if the values were comparable across all the groups and if any language configuration exerted any influence on the L1 Polish VOT durations.

As can be seen in Table 181, the VOT values for Polish /p/ were comparable for groups A, C, and D – with only group B demonstrating slightly longer durations.

Table 181. Descriptive statistics for Polish VOT for /p/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	22.8	7.9	21.3
B	L3 Ger_L2 Eng	26	28.2	9.7	30.8
C	L3 Eng_L2 Ger	33	23.3	7.2	21.5
D	L3 Eng_L2 Fre	28	22.9	8.3	19.5

Similar patterns were observed for the Polish VOT for /t/ and /k/ as presented in Tables 182 and 183.

Table 182. Descriptive statistics for Polish VOT for /t/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	29.8	7.8	28.6
B	L3 Ger_L2 Eng	26	33.9	8.2	31.1
C	L3 Eng_L2 Ger	33	31.3	7.1	30.8
D	L3 Eng_L2 Fre	28	31.5	8.7	29.7

Table 183. Descriptive statistics for the Polish VOT for /k/

Group	Languages	N	Mean	SD	Median
A	L3 Fre_L2 Eng	38	50.4	9.1	50.8
B	L3 Ger_L2 Eng	26	56.0	7.3	55.2
C	L3 Eng_L2 Ger	33	52.2	10.0	52.7
D	L3 Eng_L2 Fre	28	48.8	8.9	49.8

The across group analyses for individual plosives demonstrated no significant differences for the VOT measures in Polish apart from two exceptions: VOT durations for /p/ and /k/ proved significantly higher in group B with L3 German_L2 English than in group A with L3 French_L2 English (for /p/ Zcorr=-2.2, $p<0.05$, for /k/ t=-2.6, df=62, $p<0.05$); however, this difference was within the 5-10 ms range. Similarly, for the Polish mean VOT durations significant differences were found between groups A and B (Zcorr=-2.7, $p<0.05$) as well as B and C (Zcorr=2.4). The remaining comparisons pointed to comparable results for the Polish VOT between all the language groupings.

Recapitulating, the L1 Polish VOT values remained rather stable across the different language groups, especially if we adopt the 5 ms variability range as acceptable. It follows that L1 Polish was in general not susceptible to any regressive transfer from the participants' L2s or L3s except for somewhat higher VOT durations in the case of Group B with L3 German_L2 English which demonstrated some traces of cross-linguistic influence, i.e. the long-lag VOT of L2 English reinforced by L3 German could have exerted some influence on the L1 Polish, reflected in the overshoot of the VOT durations.

7.2. Across studies comparison

The conducted across-studies comparisons aimed to compare the global performance on L3 phonological measures investigated through the series of conducted studies. The global analysis will consist of the following components: Spearman's rank correlation analyses and mixed-effect model analyses that will be presented in the following subsections.

7.2.1. Spearman's rank correlation analysis

A global analysis was performed for all the participants of the four groups (A, B, C, D) treated jointly. The dependent variables involved selected measures of performance in the third language accounted for in the present series studies, i.e. the metaphonological awareness composite score, accentedness ratings composite score and goodness of L3 VOT fit to control values in the L3, L2 and L1. The number of dependent variables had to be limited so that they could be entered into a mixed model analysis. The independent variables included several variables, including the participants' profile data, and some linguistic factors as stipulated below.

Dependent variables:
1. Metaphonological awareness (MPhA) composite score
2. Accentedness ratings composite score
3. L3 VOT FIT_L3 – goodness of fit to control VOT in L3
4. L3 VOT FIT_L2 – goodness of fit to control VOT in L2
5. L3 VOT FIT_L1 – goodness of fit to control VOT in L1

Independent variables:
1. L2 YFT – years of formal training in L2
2. L2 AOL – beginning age of L2 learning
3. L3 YFT – years of formal training in L3
4. L3 AOL – beginning age of L3 learning
5. N-TOTAL – total number of foreign languages known
6. Stay in L3 – stay in L3 speaking countries
7. Age – participants' age
8. eval L3 comp – self-evaluation of L3 overall competence
9. eval L3 pron – self-evaluation of L3 pronunciation
10. L2 Prof – language proficiency in L2
11. L3 Prof – language proficiency in L3
12. Gender – participants' gender
13. Year – year of studies
14. L3/L1 distance – typological distance between L3 and L1 (with respect to VOT)
15. L3/L2 distance – typological distance between L3 and L2 (with respect to VOT)

16. IDENT L1 – identification of participants' L1 as correct L1 (based on accent ratings)
17. IDENT L2 – identification of participants' L1 as their respective L2
18. IDENT other – identification of participants' L1 as their other languages

The independent variables were mostly described on an ordinal scale, with the exception of gender and typological distance which were nominal (0 vs. 1). The nominal variables cannot be accounted for in the Spearman's correlation; however, they can be included in the mixed model analysis.

Table 184 presents the results of the Spearman's rank correlations between variables, with statistically significant correlation coefficients ($p<0.05$) highlighted in bold.

The results of the performed joint correlations were as follows. The metaphonological awareness (MPhA) composite score was found to correlate moderately to weakly with the following variables: the year of studies (r_s=0.46), L1 identification as respective L2 (r_s=0.39), L2 proficiency level (r_s=0.38), the participants' age (r_s=0.24), and the number of foreign languages known (r_s =0.24).

The accentedness ratings composite score was found to correlate moderately to weakly with the following variables: L3 proficiency level (r_s =0.62), self-evaluation of L3 pronunciation (r_s =0.5), self-evaluation of L3 overall competence (r_s =0.48), the years of formal training in L3 (r_s =0.44), the age of onset of L3 learning (r_s = -0.39), L2 proficiency level (r_s =0.25), and the length of stay in L3 speaking countries (r_s =0.22).

The goodness of the L3 VOT fit to the L3 control values correlated with the following variables: the goodness of the L3 VOT fit to the L1 control values (r_s =0.74), the years of formal training in L3 (r_s = -0.44), the age of onset of L2 learning (r_s = -0.41), the age of onset of L3 learning (r_s =0.39), the goodness of the L3 VOT fit to L2 control values (r_s =0.290, and age (r_s = - 0.22).

Table 184. Fig Joint Spearman's rank correlations for all participants (the correlation coefficients in bold are statistically significant at p<0.05)

Zmienna	MPhA score	Accent rating score	L3 VOT FIT_L3	L3 VOT FIT_L2	L3 VOT FIT_L1	L2_YFT	L2_ AOL	L3_YFT	L3_ AOL	N_Ln	Stay in L3	Age	eval L3 comp	eval L3 pron	L2_ Prof	L3_ Prof	Year	IDENT L1	IDENT L2
Accent rating	0.17																		
L3 VOT FIT_L3	-0.08	0.13																	
L3 VOT FIT_L2	-0.06	0.09	**0.29**																
L3 VOT FIT_L1	-0.11	0.07	**0.74**	**0.71**															
L2_YFT	-0.03	-0.02	**0.33**	**-0.19**	**0.22**														
L2_AOL	0.09	0.18	**-0.41**	**0.29**	**-0.19**	**-0.80**													
L3_YFT	0.08	**0.44**	**-0.44**	**0.29**	-0.07	**-0.22**	**0.43**												
L3_AOL	-0.07	**-0.39**	**0.39**	**-0.22**	0.09	**0.26**	**-0.31**	**-0.87**											
N_TOTAL_Ln	**0.24**	-0.08	0.15	0.01	0.09	0.13	-0.15	**-0.19**	0.13										
Stay in L3	0.02	**0.22**	0.04	0.09	0.08	-0.01	0.04	0.14	-0.11	0.16									
Age	**0.24**	0.19	**-0.22**	**0.27**	0.02	0.04	**0.42**	**0.40**	-0.15	-0.03	-0.06								
eval L3 comp	0.05	**0.48**	-0.17	**0.30**	-0.03	**-0.34**	**0.43**	**0.43**	**-0.33**	0.00	**0.22**	**0.18**							
eval L3 pron	-0.01	**0.50**	-0.05	**0.26**	0.11	-0.14	0.17	**0.38**	**-0.30**	-0.06	**0.18**	0.08	**0.68**						
L2_Prof	**0.38**	**0.25**	0.07	-0.05	0.05	**0.39**	-0.13	0.04	0.11	**0.20**	0.06	**0.42**	-0.13	-0.05					
L3_Prof	0.07	**0.62**	-0.17	**0.40**	0.04	**-0.27**	**0.45**	**0.68**	**-0.56**	-0.10	**0.20**	**0.34**	**0.70**	**0.58**	0.05				
Year	**0.46**	0.06	-0.11	0.06	0.00	0.16	0.14	**0.23**	-0.09	0.12	-0.09	**0.59**	0.03	0.00	**0.43**	0.14			
IDENT L1	-0.06	0.10	0.02	0.05	0.06	0.07	-0.15	0.03	-0.09	-0.13	0.02	-0.13	0.08	0.04	-0.03	0.02	-0.04		
IDENT L2	**0.39**	0.05	-0.15	-0.02	-0.11	-0.11	0.19	0.18	-0.16	0.01	0.04	0.20	0.09	0.03	0.16	0.14	**0.24**	**-0.46**	
IDENT other	-0.16	-0.14	0.07	-0.15	-0.09	0.03	-0.09	-0.18	0.16	0.16	0.03	-0.19	-0.14	-0.11	**-0.26**	-0.19	**-0.24**	**-0.51**	**-0.36**

As far as the across studies correlations are concerned, there were no significant relations reported between the selected measures of pronunciation performance in the third language. The only significant correlation was found to hold between the metaphonological awareness and L1 identification as the L2 (r_s =0.39). Other correlations between the MPhA composite score and the remaining L3 pronunciation performance measures proved insignificant, i.e. MPhA and accentedness ratings (r_s =0.17), MPhA and L1 identification as L1 (r_s = -0.06), MPhA and L1 identification as other languages (r_s = -0.16), MPhA and L3 VOT goodness of fit to L3 control (r_s = -0.08), MPhA and L3 VOT goodness of fit to L2 control (r_s = -0.06), and the MPhA and L3 VOT goodness of fit to L1 control (r_s = -0.11). The accentedness ratings (AR) composite score was not found to correlate significantly with any of the remaining L3 pronunciation measures under investigation, i.e. AR and L1 identification as L1 (r_s = 0.1), AR and L1 identification as L2 (r_s = 0.5), AR and L1 identification as other languages (r_s = -0.14), AR and L3 VOT goodness of fit to L3 control (r_s = 0.13), AR and the L3 VOT goodness of fit to L2 control (r_s = 0.09), and the AR and the L3 VOT goodness of fit to L1 control (r_s = 0.07). Finally, L1 identification measures did not correlate significantly with any of the VOT goodness of fit measures, with the Spearman's rank ratio being in the range of 0.02 to -0.15.

Conversely, there were statistically significant correlations reported within particular measures, i.e. within the L1 identification categories and within the L3 VOT goodness of fit categories. In the former case, there was a moderate inverse correlation between L1 identification as the L1 and L1 identification as the L2 (r_s = -0.46), between L1 identification as the L1 and L1 identification as other languages (r_s = -0.51), and between L1 identification as the L2 and L1 identification as other languages (r_s = -0.36). In the case of the goodness of fit ratings, a strong positive correlation was reported between the L3 VOT goodness of fit to L3 control and L3 VOT goodness of fit to L1 control (r_s = 0.74), and a weak one between the L3 VOT goodness of fit to L3 control and L3 VOT goodness of fit to L2 control (r_s = 0.29).

7.2.2. A mixed-effects model analysis

A mixed-effects model analysis was applied to account for both fixed and random effects in the generated joint data. As with all regression models, the purpose was to describe a response variable as a function of the predictor variables. The advantage of mixed-effects models, however, is that they recognize correlations within sample subgroups and therefore provide a compromise between ignoring data groups and fitting each group with a separate model. On the one hand, fixed effects represent population parameters that are assumed to be the same each time data is collected. Random effects, on the other hand, are sample-dependent random variables and act like additional error terms. Due to there being a large number of the independent variables only those variables that were correlated with a particular dependent variable were selected for the mixed model analysis. Another criterion for the selection was weak intercorrelations between independent variables.

The first regression analysis was performed for the metaphonological awareness (MPhA) composite score. The selected independent variables that were entered into the analysis were those that had the highest correlation coefficients with the dependent variable, i.e. in this case these were the year of studies and L1 identification as the L2. Further, the nominal variables of gender and typological distance between L3/L1 and L3/L2 were entered into the analysis. It appears that in the case of the MPhA composite score a statistically significant influence is exerted only by the year of studies. This means that mean MPhA values vary significantly as a function of the participants' different years of study. Table 185 presents the generated results.

Table 185. ANOVA results for the MPhA composite score

MPhA Composite score	Effect (F/R)	df - Effect	MS – Effect	df – error	MS – error	F	p
{1}Year	fixed	3	1743.13	24.91	172.68	10.09	**0.000155**
{2}Gender	fixed	1	0.63	30.64	150.36	0.00	0.948949
{3}L3/L1 distance	fixed	1	91.58	8.07	244.53	0.37	0.557408
{4}L3/L2distance	fixed	1	155.78	11.19	217.55	0.72	0.415165
{5}IDENT L2	random	6	279.13	25.00	138.97	2.01	0.102341

$p<0.05$

In a subsequent analysis, the dependent variable of accentedness rating (AR) composite score was correlated with the independent variables that displayed the highest correlations, i.e. the self-evaluation of L3 competence, the self-evaluation of L3 pronunciation, and the L3 proficiency level. The accentedness ratings composite score was calculated as a means of the three rating parameters including foreign accentedness, intelligibility and perceived pronunciation accuracy. The nominal variables of gender and typological distance between L3/L1 and L3/L2 were entered into the analysis.

The analysis demonstrated that such variables as the self-evaluation of L3 competence, the self-evaluation of L3 pronunciation as well as the L3 proficiency level exert a statistically significant influence on the dependent variable of accentedness ratings. A statistically significant interaction was also found between the self-evaluation of L3 pronunciation and gender, as well as between the self-evaluation of L3 competence and the L3/L1 typological distance, which means that the interaction of these independent variables also has an effect on the accentedness ratings score. The results of the mixed model analysis for accentedness ratings are presented in Table 186.

Table 186. ANOVA results for accentedness ratings composite score

Accentedness ratings score	Effect (F/R)	df – Effect	MS – Effect	df – error	MS – error	F	p
{1}eval L3 comp	fixed	3	3.45	23.00	0.36	9.49	**0.000287***
{2}eval L3 pron	fixed	3	1.90	23.00	0.36	5.24	**0.006675***
{3}L3_Prof	fixed	4	2.49	23.00	0.36	6.85	**0.000874***
{4}Gender	fixed	1	0.73	23.00	0.36	2.01	0.169717
{5}L3/L1 distance	fixed	1	0.00	23.00	0.36	0.01	0.916939
{6}L3/L2distance	fixed	1	0.17	23.00	0.36	0.47	0.499594
1*2	fixed	4	0.22	23.00	0.36	0.60	0.663272
1*3	fixed	3	0.71	23.00	0.36	1.96	0.147757
1*4	fixed	1	0.05	23.00	0.36	0.14	0.709912
1*5	fixed	2	2.10	23.00	0.36	5.78	**0.009240***
1*6	fixed	2	0.01	23.00	0.36	0.02	0.976528
2*3	fixed	5	0.49	23.00	0.36	1.36	0.277057
2*4	fixed	1	2.43	23.00	0.36	6.68	**0.016549***
2*5	fixed	2	0.17	23.00	0.36	0.47	0.633734
2*6	fixed	0	0.00				
3*4	fixed	3	0.39	23.00	0.36	1.06	0.384952
3*5	fixed	0	0.00				
3*6	fixed	1	0.02	23.00	0.36	0.05	0.831430
4*5	fixed	1	0.58	23.00	0.36	1.60	0.218483
4*6	fixed	0	0.00				
5*6	fixed	0	0.00				

* $p<0.05$

In the following mixed-effects model analysis, the dependent variable of the L3 VOT goodness of fit to the control VOT for the L3 was correlated with the independent variables that demonstrated the highest correlations, i.e. the years of formal training in L2 (L2 YFT), the age of onset of L2 learning (L2 AOL), the years of formal training in L3 (L3 YFT), the age of onset of L3 learning (L3 AOL) and the participants' age. As in the previous cases, the nominal variables of gender and typological distance between L3/L1 and L3/L2 were part of the analysis.

The results show that only the L3/L1 typological distance exerts a statistically significant influence on the dependent variable. It appears that the mean values of the goodness of fit between L3 VOT and the control VOT for the L3 differ as the function of typological distance between the L3 and the L1. Detailed results of the analysis are presented in Table 187.

Table 187. ANOVA results for L3 VOT goodness of fit to L3 control

L3 VOT FIT to control L3	Effect (F/R)	df – Effect	MS – Effect	df – error	MS – error	F	p
{1}Gender	fixed	1	1.04	85.71	4.33	0.24	0.625532
{2}L3/L1 distance	fixed	1	572.06	9.54	3.34	171.40	**0.000000***
{3}L3/L2distance	fixed	1	7.48	8.56	3.90	1.92	0.201176
{4}L2_YFT	random	14	3.22	56.62	4.49	0.72	0.748093
{5}L2_AOL	random	14	5.29	84.95	4.14	1.28	0.238805
{6}L3_YFT	random	15	4.45	44.65	4.12	1.08	0.401165
{7}L3_AOL	random	15	3.88	54.50	4.36	0.89	0.578307
{8}Age	random	6	1.93	50.00	4.71	0.41	0.868515

*p<0.05

In yet another analysis, the dependent variable of the L3 VOT goodness of fit to the control VOT for the L2 was correlated with the independent variables that demonstrated the highest correlations with this variable, such as the years of formal training in L2 (L2 YFT), the age of onset of L2 learning (L2 AOL), the years of formal training in L3 (L3 YFT), the age of onset of L3 learning (L3 AOL), age, the self-evaluation of L3 competence, the self-evaluation of L3 pronunciation and the L3 proficiency level.

The results demonstrated that the dependent variable of L3 VOT goodness of fit to the control VOT for the L2 was significantly influenced by the self-evaluation of L3 competence, the L3 proficiency level, the age of learning of L3 and the typological distance between L3 and L2 (see Table 188).

Table 188. ANOVA results for L3 VOT goodness of fit to L2 control

L3 VOT FIT to control L2	Effect (F/R)	df – Effect	MS – Effect	df – error	MS – error	F	p
{1}eval L3 comp	fixed	4	50.61	24.36	6.02	8.41	**0.000208***
{2}eval L3 pron	fixed	3	5.14	45.79	7.43	0.69	0.561742
{3}L3_Prof	fixed	4	27.11	33.52	8.78	3.09	**0.028773***
{4}Gender	fixed	1	3.66	22.76	10.94	0.33	0.568424
{5}L3/L1 distance	fixed	1	0.19	4.12	8.16	0.02	0.885899
{6}L3/L2distance	fixed	1	85.89	22.25	12.75	6.74	**0.016410***
{7}L2_YFT	random	14	3.77	28.49	10.05	0.38	0.971666
{8}L2_AOL	random	13	11.96	22.53	7.26	1.65	0.144561
{9}L3_YFT	random	15	3.71	14.28	13.41	0.28	0.990758
{10}L3_AOL	random	14	13.50	10.67	2.96	4.55	**0.008666***
{11}Age	random	6	3.21	7.00	2.53	1.27	0.378597

*p<0.05

Finally, the dependent variable of the L3 VOT goodness of fit to the control VOT for the L1 was correlated with the independent variables that exhibited the highest correlations, i.e. the years of formal training in L2 (L2 YFT), and the age of onset of L2 learning (L2 AOL). The results of the mixed model for the L3 VOT goodness of fit to the control L1 VOT did not demonstrate any statistically significant results for any of the examined interactions (see Table 189).

Table 189. ANOVA results for L3 VOT goodness of fit to L1 control

L3 VOT FIT to control L1	Effect (F/R)	df – Effect	MS – Effect	df – error	MS – error	F	p
{1}Gender	fixed	1	32.01	3.84	9.64	3.32	0.145553
{2}L3/L1 distance	fixed	1	31.29	5.60	22.60	1.38	0.286953
{3}L3/L2distance	fixed	1	7.48	3.48	7.54	0.99	0.383264
{4}L2_AOL	random	15	9.69	2.78	16.56	0.59	0.790509
1*2	fixed	1	25.86	101.18	12.18	2.12	0.148191
1*3	fixed	1	2.46	16.61	10.85	0.23	0.640444
1*4	random	7	10.32	96.24	12.28	0.84	0.556933
2*3	fixed	0	0.00				
2*4	random	8	19.28	0.38	2.50	7.72	0.525578
3*4	random	2	4.53	87.00	12.26	0.37	0.692272

*p<0.05

7.2.3. Summary and conclusions

Summing up, the performed global analysis for all the participants treated jointly demonstrated significant patterns of correlations between the selected measures of performance in the third language and some linguistic factors and factors regarding the participants' background. The accentedness ratings score correlated significantly with the self-evaluation of L3 competence, the self-evaluation of L3 pronunciation as well as the L3 proficiency level. For the metaphonological awareness composite score, a statistically significant influence was exerted only by the variable of the year of studies. In the case of the goodness of fit between L3 VOT and native control, the mean values differed as the function of typological distance between the L3 and the L1. The highest number of correlations were reported for the L3 VOT goodness of fit to the L2 control, including the self-evaluation of L3 competence, the L3 proficiency level, the age of learning of L3 and the typological distance between the L3 and the L2.

Nevertheless, the across-studies analyses did not point to any significant correlations between the selected measures of pronunciation performance in the third language, the only exception being a moderate correlation between the metaphonological awareness and L1 identification as the L2. This may indicate that the nature of these separate measures was very different, ranging from a global perceptual assessment of foreign accentedness, through a local specific phonetic measure of voice onset time, to a quantification of verbal protocols of metalinguistic awareness. The inherent diversity of the applied measures and tasks could therefore explain the lack of correlations across the studies.

7.3 Research hypotheses verification

In the present section, the research hypotheses stipulated in Chapter 2 will be tested based on the findings of the conducted studies. The major goal of the performed series of studies was to investigate which factors condition, and to what extent, cross-linguistic influence (CLI) in the acquisition of third language phonology. The explored factors included: the L1 effect, the L2 status, the typological proximity or the combined effect of all these factors.

In order to verify the research hypotheses put forward previously, four comparisons were performed between the study groups (A, B, C, and D) with a view to investigating whether any of these factors determine the results, and if so, which ones. Based on the previously presented detailed results of the three studies (see chapters 4-6), the additional comparative analyses across the studies and across the groups were now subject to further interpretation. The performed comparisons involved six parameters selected from the three conducted series of studies:

Study I – global perception of foreign accentedness:

(1) foreign accent composite score (calculated as the mean value of the three subcomponents, i.e. accentedness, comprehensibility and pronunciation correctness)

(2) L1 identification patterns (as the L1, L2 or other languages).

Study II – acoustic measures of a local phonetic feature of voice onset time:

(3) L3 VOT goodness of fit to the control values

(4) VOT values in the L1, L2, L3

Study III – quantitative and qualitative analysis of metacognitive performance:

(5) metaphonological awareness (MPhA) composite score

(6) MPhA metacomments.

In line with the previously stated aim, a number of comparative analyses were performed for the following four groupings:

1) A vs. B – a group pairing with the same L1 and L2, but different L3s

A: L1 Polish, L2 English, L3 French vs. B: L1 Polish, L2 English, L3 German

2) C vs. D – a group pairing with the same L1 and L3, but different L2s

C: L1 Polish, L2 German, L3 English vs. D: L1 Polish, L2 French, L3 English

3) B vs. C and A vs. D – typologically related and unrelated pairs of groups

B: L1 Polish, L2 English, L3 German vs. C: L1 Polish, L2 German, L3 English

A: L1 Polish, L2 English, L3 French vs. D: L1 Polish, L2 French, L3 English

4) A vs. B and C vs. D – pairs of groups with the same L2 and different L2s

A: L1 Polish, L2 English, L3 French vs. B: L1 Polish, L2 English, L3 German
C: L1 Polish, L2 German, L3 English vs. D: L1 Polish, L2 French, L3 English

The comparative analysis of the selected L3 phonetic performance measures in particular groupings was designed to verify the research questions and hypotheses stipulated earlier in Chapter 2 section 2.5.1.

7.3.1. Foreign accentedness composite score

The research hypothesis verification first concerned the group analyses performed for the global perception of foreign accentedness. The foreign accent composite score was calculated as the mean value of the three subcomponents, i.e. accentedness, comprehensibility and pronunciation correctness.

1) A vs. B – the same L1 and L2, but different L3s

The group comparison between A and B (the same L1 and L2 but different L3s) did not show any significant difference as far as the foreign accentedness scores are concerned. Since the average rating scores were comparable for both groups irrespective of their various L3s (French and German), the findings may be interpreted as pointing to the relevance of the L2 status or L1 effect for global accentedness, as these two variables were constant in both groups, resulting in traces of L1- or L2-accentedness. Conversely, the results may not be attributed to the impact of L3 features as the accentedness scores in groups A and B were comparable irrespective of the different L3s.

2) C vs. D – the same L3 and L1, but different L2s

The results of the comparison between groups C and D with the same L3 but different L2s indicate statistically significant differences in the foreign accentedness ratings composite score. Group D (L3 English_L2 French) outperformed group C (L3 English_L2 German) in the rating scores. Since the performance in L3 English was not comparable between the groups for this parameter, it may be interpreted as the L2 exerting an impact on the L3 phonological performance. Further, as it was Group D with a typologically more distant set of L2 and L3 (French and English) that received better accentedness scores than the more typologically related Group C (with L2 German and L3 English), then these results may lead us to question typology as a facilitating factor in L3 phonological acquisition with respect to accentedness. An alternative explanation might be that German and English are not perceived as typologically similar in terms of global accentedness due to different articulatory settings, thus we may not expect any facilitation in this respect.

3) Comparison between typologically related groups and unrelated groups

As far as the comparison between typologically related and unrelated groups is concerned, the differences in the foreign accentedness scores were not found to be statistically significant either for the typologically related groups (B vs. C) or the typologically unrelated ones (A vs. D). The two pairs of groups scored similarly on foreign accentedness composite ratings, which may be interpreted as counterevidence against the relevance of typology as a determiner of L3 phonological performance. It appears that the presence or absence of typological proximity between particular language pairs (L2 and L3) did not exert any major impact on the global perception of accentedness. Moreover, since there were no major differences in the accentedness scores between both pairs of groups, one may not rule out the underlying L1 effect on the L3 phonological performance as it was the first language that they all had in common.

4) Comparison between groups with the same L2 and different L2s

The comparison of the foreign accentedness scores of pairs of groups with the same L2 (A vs. B) and with different L2s (C vs. D) demonstrated no significant differences in the former group, yet statistically different accent rating scores in the latter. Since the groups with the same L2 performed similarly, whereas the other two groups did not, it may be attributed to an impact of the L2. It seems that keeping the L2 constant for groups A and B resulted in comparable accentedness scores in the L3 performance, although the L3s differed. On the other hand, the L3 English was rated as differently accented in groups C and D, possibly due to the difference in the L2s. These results may thus be interpreted as verifying the importance of CLI from the L2 for the perception of global accentedness.

7.3.2. L1 identification patterns

The L1 identification task was performed on L3 samples as part of the foreign accentedness ratings. The raters were asked to identify the mother tongue of the speaker on the basis of the samples of the L3 speech. The findings demonstrated different patterns of L1 identification – with L1 identification as the correct L1, as the L2 or as other languages. The respective group comparisons and the resulting hypotheses verification shall be presented in the following subsections.

7.3.2.1. L1 identification as L1

As far as the correct L1 identifications (i.e. L1 identification as L1 Polish) are concerned, the percentage scores were in the range of 30-40%; the lowest was for Group C with L3 English_L2 German (29% identification as Polish); intermediate for Group A with L3 French_L2 English (36%) and Group D with L3 English_L2 French (37.8%); and highest for Group B with L3 German_L2 English (41.6%). Between group comparisons yielded relatively few statistically significant differences, which will be discussed below.

1) A vs. B – the same L1 and L2, but different L3s

The comparative analysis between Groups A and B with the same L2 (English) but different L3s (French and German) failed to indicate any statistically significant differences. The performance of different L3 groups did not vary with respect to L1 identification correctly as L1, thus the findings may be attributed to the impact of the L2 status or the L1 effect since the L1 and L2 were kept constant in both groups. Further, it appears that the L3 characteristics did not influence the L1 identification ratings.

2) C vs. D – the same L3 and L1, but different L2s

In the comparison between Groups C and D with the same L3 (English) but different L2s (German and French), no significant differences were found in the L1 identification as L1 patterns. Irrespective of the different L2s, the groups received similar L1 identification scores, and therefore it may be interpreted as a counterevidence against any significant impact of the L2 on the L3 phonological performance for this parameter.

3) Comparison between typologically related and unrelated pairs of groups

As far as the comparison between typologically related and unrelated pairs of groups is concerned, the former pair (B vs. C) exhibited significant differences in their L1 identification as L1 scores, but the latter (A vs. D) did not. We could have expected a similar performance in the L1 identification in the typologically related pair of groups, yet this expectation was not substantiated by the data. Group B with L3 German and L2 English had significantly higher L1 identification as L1 scores than Group C with L3 English_L2 German in spite of the typological closeness between the respective L2 and L3. It appears that the typological closeness between the L2 and L3 was not a determining factor contributing to the perceived cross-linguistic influence as reflected by the L1 identification percentages since the scores were not comparable for Groups B and C. Conversely, the percentages of the L1 identification as L1 were similar in the case of the less typologically related group pair A and D (with the French and English mirror design), which confirms the assumption that typology should not necessarily be regarded as a conditioning factor in determining cross-linguistic influence in this respect.

4) Comparison between pairs of groups with the same L2 and different L2s

The comparison of the percentage of correct L1 identifications as L1 between pairs of groups with the same L2 (A vs. B) and different L2s (C vs. D) did not generate any statistically significant differences, thus disconfirming an assumption that the L2 may be regarded as a predictor of the L3 performance with respect to first language identifications based on L3 samples.

On the whole, in the accent rating task in the third language, 30-40% of the samples were correctly identified as having Polish as their L1. The correct L1 identification was stronger than the L1 identification as the L2 or other languages treated as separate categories, and therefore the L1 status comes to the fore as an influencing factor in this respect. However, it did not prove to be overwhelmingly dominant, thus questioning the status of the first language as the only, or basic, constraint in third or additional language acquisition.

7.3.2.2. L1 identification as L2

As far as identification of the native tongue of the L3 speakers as the respective L2 is concerned, there were no statistically significant differences in the percentage scores between the groups or pairs of groups. The L2 identification percentages were relatively low, with the mean being the lowest for Group A with L3 French_L2 English (8% identification as English), intermediate for Group B with L3 German_L2 English (10.6% identification as German) as well as Group D with L3 English_L2 French (14% identification as French), and the highest for Group C with L3 English_L2 German (17.6% identification as German).

Since there were no significant differences in any of the between group comparisons, no conclusions can be drawn with respect to the factors that determined the identification of the first language of the L3 speakers as their respective L2s. The low L1-identification-as-L2 scores indicate, however, that cross-linguistic influence from the second language was not particularly salient in the L3 speech. On the whole, this finding provides evidence that the L2 was found to be only weakly influential in determining the L1 identification patterns in third language accentedness ratings. However, there were visible traces of L2-accentedness and, actually, it was always the respective L2 that had the second highest identification score following the L1 in particular groups. Conversely, it is the L1 effect that proved to have more impact as a factor determining cross-linguistic influence in the L3 speech performance since the L1-identification-as-L1 scores based on L3 samples were considerably higher than the L1 identification as the L2.

7.3.2.3. L1 identification as other languages

The identification of the native tongue of the L3 speakers as Ln treated jointly, i.e. languages other than their respective L1 or L2, yielded composite results in the range of 48-56%, which is the highest percentage compared to the L1 identifications as the L1 (30-40%) or as the L2 (8-17%).

Only one of the between group comparative analyses proved statistically significant, namely the one for groups with the same L2 but a different L3. The mean values for L1 identification as Ln were significantly higher for Group A with L3 French_L2 English than Group B with L3 German_L2 English. Since the performance between different L3 groups varied irrespective of the fact that they shared the same L2, the results may not be attributed to CLI from the L2, but rather other factors such as various L3 characteristics.

The comparison between Groups C vs. D with the same L3 and L1, but different L2s did not yield any significant differences. The percentage of L1 identification as Ln was comparable irrespective of the different L2s, which would not provide support for CLI from the L2. Further, no significant difference in performance was noted for the typologically related (B vs. C) vs. unrelated pairs of groups (A vs. D), which disconfirms any impact of the typological distance between the L2 and L3 with respect to L1 identification as Ln.

7.3.3. L3 VOT goodness of fit to the control values

Study II involved the acoustic measures of a local phonetic feature of voice onset time. Two parameters that were subject to comparative analyses across groups and hypotheses verification included (1) the VOT goodness of fit to the control values for the L3 and L2, (2) the VOT duration measures in the L1, L2, L3 for particular languages. The following sections (7.1.3 and 7.1.4) will discuss the two parameters separately.

7.3.3.1. L3 VOT goodness of fit to L3 controls

1) A vs. B – the same L1 and L2, but different L3s

When the L3 VOT goodness of fit scores were juxtaposed for Groups A and B, which were characterized by the same L2 but different L3s, the analysis pointed to statistically significant differences. In Group A, the L3 French VOT values were by 25% too high compared to the French control VOT, possibly due to there being high L2 English VOT values which would partially support CLI from the L2 on the L3 VOT durations. On the other hand, in Group B, the L3 German VOT values appeared too low by approximately 20% compared to the German control VOT, resulting in varying patterns of goodness of fit for the L3 VOT values in the two groups. Irrespective of the fact that they shared the same L2, i.e. English, the performance of the groups with different L3 varied, thus indicating that the findings may not be so much attributed to CLI from the L2 but rather to the impact of different L3 characteristics and possibly an underlying L1 effect for group B but not A.

2) C vs. D – the same L3, but different L2s

The comparison of the L3 VOT goodness of fit to L3 control VOT in Groups C vs. D with the same L3 but different L2s, did not show any statistically significant differences. In both groups the L3 English VOT scores were approximately 30% lower than the English control VOT values. It appears that the goodness of fit of the L3 English VOT to the English control values was comparable in groups C and D irrespective of the L2. The findings thus demonstrate that the L2 exerted no or little impact on the L3 performance in terms of the VOT values, but rather that the underlying L1 effect was visible, resulting in shorter lag VOT durations. Otherwise, we would expect better goodness of fit scores in Group C (due to the higher VOT durations in L2 German) than in Group D, where the scores should be worse due to there being shorter L2 French VOT durations.

3) Comparison between typologically related groups and unrelated groups

The comparison between the typologically related and unrelated groups in terms of their L3 VOT goodness of fit yielded significantly different values in both pairs of groups. In the typologically related groups (B vs. C) the difference was smaller and both groups followed a similar pattern, whereas as in the typologically unrelated pair (A vs. D) the difference was larger and the groups demonstrated an opposite pattern. In the former pair, the VOT values in L3 English in Group C were by 30% lower than the English control VOT values, whereas in Group D the L3 German VOT was lower by 20% than the German VOT controls. On the other hand, in the latter pair, in Group D the L3 English values were lower by approximately 30% compared to the English controls, yet the L3 French VOT was higher by 25% than the French control values in Group A. The typologically related pair (B & C) behaved somewhat more similarly and demonstrated more comparable VOT goodness of fit patterns than the typologically less related pair (A & D), which may be interpreted as partial or weak evidence for the facilitating effect of typology with respect to the acquisition of the L3 VOT patterns. It appears that typological proximity may be more important for the acquisition of a local feature such as VOT in the third language.

4) Comparison between groups with the same L2 and different L2s

The final comparative analysis between the pairs of groups with either the same L2 (A vs. B) and different L2s (C vs. D) demonstrated significant differences in the L3 VOT goodness of fit scores for the former pair and not in the latter. Irrespective of the same L2, i.e. English in Groups A and B, the L3 VOT scores followed different goodness of fit patterns in the respective L3s (French and German). In the case of L3 French this may be a partial influence of the high L2 English VOT values; however, this effect does not hold true for L3 German. Further, had typology been a determining factor, we might expect better goodness of fit scores in group B. On the other hand, in Groups C and D, the L3 English VOT demonstrated comparable goodness of fit values irrespective of the varying L2s (French and German). This result may be interpreted as an item of counterevidence against CLI from the L2 as we would expect varying results for Groups C and D – which was actually not the case.

7.3.3.2. L2 VOT goodness of fit to L2 controls

As far as the goodness of fit of L2 VOT to the control values in the respective languages is concerned, the majority of the groups demonstrated a much better match than in the case of the L3 values, which fell beyond the target VOT values. Interestingly, for L2 English in Groups A and B as well as for L2 French in Group D, there was only a discrepancy of 2 to 6% from the control VOT values in the respective languages. In the case of L2 German in Group C, the VOT measurements were by 30% lower than the control values. The German control VOT durations, however, exceeded considerably the VOT values reported in the literature for German, which may to some extent weaken their full validity and representativeness. All in all, the L2 VOT values approximated the target values much better than the L3 VOT durations, which validates the greater stability of the phonological systems of the second languages compared to that of the third languages.

1) A vs. B – the same L1 and L2, but different L3s

A comparative analysis of groups A and B sharing the same L2 but differing with respect to the L3s demonstrated no significant differences in the L2 VOT goodness of fit ratings. In both groups, the L2 English values approximated quite successfully the target ones (below 6% deviation from the control values) regardless of the different L3s (French and German) or different typological distance between the L2 and L3. These findings may be interpreted as a counterevidence against any regressive cross-linguistic influence from the third language to the better established, and thus more stable, VOT values in the second language. Summing up, the more stable system of the L2 was not found to be susceptible to a later acquired L3.

2) C vs. D – the same L3 and L1, but different L2s

The results of the C vs. D group comparison, which involved the same L3 but different L2s, yielded statistically significant differences. Group C displayed considerably lower L2 German VOT goodness of fit to control values than in the case of Group D with L2 French. In the latter group, a facilitative L1 effect could be at play, since Polish and French VOT durations are quite comparable (i.e. short lag) as opposed to the long lag durations in German or English. It follows from these findings that groups with different L2s behaved independently with respect to approximating the target VOT values in the L2, irrespective of their shared L3 (i.e. English). Again, no regressive influence of the L3 on the L2 was attested.

3) Comparison between typologically related groups and unrelated pairs of groups.

As far as the comparison between typologically related and unrelated pairs of groups is concerned, statistical differences in the L2 VOT goodness of fit to controls were observed in the former pair, i.e. B vs. C (with a mirrored German and English design) but not in the latter, i.e. A vs. D (with a mirrored French and English design). It may follow that the typological distance between the L2 and L3 was not found to be a factor influencing the approximation of the L2 VOT values to the target norms. Consequently, the L3 does not affect phonological performance in the L2.

4) Comparison between pairs of groups with the same L2 and different L2s

A comparative analysis between the pairs of groups with the same L2 (A vs. B) and different L2s (C vs. D) demonstrated that the former have comparable results with regard to the L2 VOT goodness of fit to controls, whereas the latter vary significantly between one another. As could have been expected, the L2 VOT patterns were determined by the acoustic characteristics of the target language.

7.3.4. VOT values for particular languages

Across the groups comparisons were performed for the VOT values separately for /p/, /t/, /k/ and then jointly for the mean VOT for all the languages involved, namely English, French, German and Polish, irrespective of their language status (L1, L2, L3). The results are discussed with respect to the verification of the hypotheses in different language sections.

7.3.4.1. VOT for English

In the case of English, the VOT analyses for separate plosives followed identical patterns, and therefore only a joint analysis for the mean VOT will be discussed in detail. The mean VOT English durations did not differ significantly ($p>0.05$) for groups with English as their L2 or groups with English as the L3. In the case of L2 English, the difference between groups with L3 French_L2 English (M=76.4) and L3 German_L2 English (M=73.5) was not found to be statistically significant (Student t-test: $t=0.59$, $p>0.05$). Likewise, the comparison of the L3 English groups i.e. with L3 English_L2 German (M=46.8) and L3 English_L2 French (M=46.2), did not yield any significantly different results (Man-Whitney test: Z corr=0.68, $p>0.05$). It follows that having different L2s (French and German) did not exert any significant influence on the L3 English mean VOT values. However, the comparisons between the VOT durations in English as the L2 and as L3 resulted in significant differences. The performed Cochran-Cox test indicated a statistically significant difference between the mean VOT durations for the L3 German_L2 English vs. the L3 English_L2 German group (tcc=5.8, $p<0.05$) with the VOT values for L2 English (M=73.5) being significantly higher than for L3 English (M=46.8). The comparative analysis between L3 English_L2 French vs. L3 French_L2 English groups yielded similar results (Man-Whitney test; Zcorr=-5.7, $p<0.05$) with the mean VOT durations in the L2 English (M=76.4) being significantly longer than in the L3 English (M=46.2).

1) A vs. B – the same L1 and L2, but different L3s

Moreover, the comparison of mean VOT values between Groups A and B with the same L2 English but varying L3s did not demonstrate any statistically different results either. In both groups the VOT durations were comparable, thus disconfirming any potential regressive transfer from the L3 to L2 since we would expect varying results of such an influence due to different VOT target values in the respective L3s (French and German).

2) C vs. D – the same L3 and L1, but different L2s

Recapitulating, the comparison between groups with the same L3 but different L2s (C vs. D) did not result in any significant differences. A conclusion can be drawn that CLI from L2 was not the factor determining L3 values, as the influence of different L2s (with French and German varying in terms of VOT durations) should be reflected in different VOT values for L3 English in Groups C and D, which was not the case.

3) Comparison between typologically related and unrelated pairs of groups

As far as the comparison between the pairs of typologically related and unrelated groups with respect to the English mean VOT is concerned, the findings indicated significantly different results for both pairs (B vs. C and A vs. D). The major factor seems to be the status of English either as the L2 or as the L3, as in the former case the mean VOT values in L2 English were higher (both in A and B) and approximated the target English durations, whereas in the latter, VOT in L3 English was significantly shorter (both in C and D) irrespective of the typological relatedness between the pairs of groups. The findings thus disconfirm any significant role of typological distance in shaping the VOT values as the performance of Groups B and C did not differ from that of Groups A and D.

4) Comparison between pairs of groups with the same L2 and different L2s

Interestingly, the comparison between pairs of groups with the same L2 (A vs. B) and with different L2s (C vs. D) was not found to be significantly different with respect to the mean VOT in English. This may be interpreted as failing to provide verification for CLI from the L2.

7.3.4.2. VOT for French

In the case of the French VOT values, a comparative analysis could involve only two groups which have French in their language repertoires, i.e. Group D with L3 English_L2 French, and Group A with L3 French_L2 English. As in the previous analysis for English, in the case of French, the analysis of the VOT values for /p, t, k/ followed identical patterns to the joint analysis for the mean VOT. The results of the A vs. D group comparison demonstrated a statistically significant difference, which may be related to the varying status of French either as the L2 (Group D) or as the L3 (Group A). In the case of L2 French, the VOT durations were shorter and approximate more the target French short lag values. Interestingly, the L3 French VOT values exceed the target ones, and this finding could be interpreted as an evidence of the influence of longer lag VOT values in L2 English on L3 French, thus indicating CLI from the L2.

7.3.4.3. VOT for German

The analysis of the German VOT in Groups B (L3 German_L2 English) and C (L3 English_L2 German) showed no significant differences for /p/, /k/ or a joint analysis of German VOT, although in all these cases the VOT values for the L3 German were slightly higher than for the L2 German. The only statistically significant differences between both groups were found for German /t/, for which the VOT duration in L3 German is significantly higher than in L2 German. These findings could be interpreted as a weak support for CLI from the L2 as some variation in the VOT duration between Group B and C could be interpreted as an influence of the long lag VOT values from L2 English on L3 German, although this influence does not seem to be as consistent and regular as in the case of the French VOT.

7.3.4.4. VOT for Polish

In L1 Polish, the comparative group analyses for individual plosives demonstrated no significant difference for the VOT measures in the Polish apart from two exceptions, i.e. the VOT durations for /p/ and /k/ proved significantly higher in Group B with L3 German_L2 English than in Group A with L3 French_L2 English; however, this difference was in the 5-10 ms range. Similarly, for the Polish mean VOT durations significant differences were found between Group B with L3 German_L2 English and Group A with L3 French_L2 English, as well as between Group B with L3 German_L2 English and Group C with L3 English_L2 German. The remaining comparisons pointed to comparable results for the Polish VOT between all the language groupings.

Recapitulating, the L1 Polish VOT values remained rather stable across different language groups, especially if we adopt the 5 ms variability range as acceptable. It follows that L1 Polish was in general not susceptible to regressive transfer from the participants' L2s or L3s except for the somewhat higher VOT durations in the case of Group B with L3 German_L2 English which demonstrated some traces of cross-linguistic influence, i.e. the long-lag VOT of L2 English reinforced by L3 German could have exerted some influence on the L1 Polish reflected in the overshoot of VOT durations. The findings grant some weak support for the potential CLI even in the case of L1 VOT values.

7.3.5. Metaphonological awareness composite score

In the case of the MPhA composite score, the group results were interpreted in the following manner.

1) A vs. B – the same L1 and L2, but different L3

The performance on MPhA between different L3 groups was found to differ significantly irrespective of the shared L2. Group B (L3 German_L2 English) outperformed group A (L3 French_L2 English) on the measure of metaphonological awareness. Since the performance of groups A and B was not the same, yet they shared the first and the second languages, the results could not be attributed to influence from either the L1 or L2, and thus other factors could have been at play. For instance, it could be the result of the impact of different properties of the respective L3s. It may also be hypothesized to be due to the influence of typology as the L2 was kept constant. The MPhA results of Group B, with closer typological links between L3 English and L3 German, were better than in Group A, with a bigger typological distance between L3 English and L2 French. Thus the typological relatedness could have reinforced the metacognitive performance scores.

2) C vs. D – the same L1 and L3, but different L2s

The metaphonological awareness in L3 English was not found to be statistically different in the C and D groups with different L2s (L2 French and L2 German), i.e. it was comparable irrespective of the L2. It may follow that the L2 status did not determine or differentiate the L3 phonological performance in terms of MPhA. An alternative interpretation is also possible, namely that the L2 learning experience exerted an impact on the metacognitive competence in the L3, irrespective of what this language actually was. Moreover, the same L3 properties could have influenced the actual score in both groups as they shared English as their L3.

3) Comparison between typologically related groups and unrelated groups

The comparison between the MPhA scores in Group B vs. C demonstrated a statistically significant difference, although these were L2-L3 mirror groups with typologically related L2s and L3s (i.e. English and German). Therefore, the impact of typological closeness does not appear to be a determining factor in this respect, as we could have expected reciprocal facilitation in both groups. In turn, the remaining two mirror groups, i.e. A vs. D were not found to be statistically different in terms of the MPhA scores. In spite of a greater typological distance between the respective L2s and L3s (i.e. English and French) these two groups scored similarly on the MPhA, which may again be interpreted as counterevidence against the relevance of typology as a determiner of L3 phonological awareness. Typology was not found to be a significant predictor for metaphonological performance in the L3. Since there were differences between all the four groups, one would need to exclude the assumption that it is only the L1 effect that influences the L3 phonological performance in terms of the MPhA.

4) Comparison between groups with the same L2 and different L2

A further between groups comparison was performed involving the groups with the same L2 (A and B) vs. the groups with different L2s (C and D). The results demonstrate that the same L2 groups were found to score significantly differently on the MPhA test (Group B outperformed A), whereas there were no statistical differences in the scores for the groups with different L2s but the same L3 (Group C performed similarly to D). It may be interpreted that it is the properties of the third language that determined the relationship, namely, the same L3 English resulted in a similar performance on the MPhA, whereas different L3s led to a significantly different performance on the MPhA. Moreover, since the first two groups did not behave in the same way whereas the remaining two groups did, the findings may not be attributed to the L2 and it does not seem to be a determining factor as far as the metaphonological performance in the L3 is concerned.

7.3.6. Metaphonological self-reports

Another part of Study III consisted in the qualitative analysis of metacomments generated in the oral protocols. Of particular significance in this analysis were self-reports of cross-linguistic influence (see 6.3.2.1 for a detailed discussion). These reports demonstrated perceived interactions between different language systems of the multilingual participants as well as the declared sources, directionality and strength of the cross-linguistic influence. The vast majority of the participants declared their second language (L2) as the main source of cross-linguistic influence in the L3 phonological acquisition, thus corroborating, to a large extent, the hypothesis of the CLI from the L2 (Group A – 48%, B – 62.5%, C – 86%, D – 50%). Furthermore, the strength of the reported L2 influence was found to be considerably higher in more typologically related groups with L2/L3 English and German (i.e. groups C and D) than in the more typologically distant English and French group pairs (i.e. A and D). A much smaller percentage of the participants indicated their first language, i.e. L1 Polish, as the main source of CLI in L3 phonological acquisition (Group A – 22%, B – 12.5%, C – 10%, D – 23%), providing only weak support for the traditionally held assumption of the greatest influence form the L1. Finally, there were considerable indications of a combined cross-linguistic influence, i.e. both the L1 and L2 were recognized as exerting an influence on the L3 phonology (Group A – 30%, B – 25%, C – 4%, D – 27%).

7.3.7. Discussion

Table 190 presents a summary of the hypotheses verification based on the comparative analysis of four grouping effects (A vs. B; C vs. D; B vs. C and A vs. D; A vs. B and C vs. D) for six measures (i.e. two measures per three studies). The effects of the L1, the L2, the L3 and typological proximity are evaluated as independent variables exerting influence (‘yes’), the lack of influence (‘no’) or some degree of CLI (‘weakly/partially’) on the phonological performance on particular measures.

Table 190. Hypotheses verification – summary

Study	Task	Groupings	CLI from L1	CLI from L2	Typology	L3 features
I.	Foreign accent ratings	A vs. B	–	–	no	no
		C vs. D	–	yes	no	–
		B/C vs. A/D	yes	–	no	–
		A/B vs. C/D	–	yes	–	–
	L1 identification as L1	A vs. B	yes	yes	no	no
		C vs. D	yes	no	no	–
		B/C vs. A/D	–	–	no	–
		A/B vs. C/D	–	no	–	–
	L1 identification as L2	All	yes	weakly yes	–	–
	L1 identification as Ln	All	–	–	–	yes
II.	L3 VOT goodness of fit	A vs. B	partially yes	weakly yes	–	yes
		C vs. D	yes	weakly no	no	yes
		B/C vs. A/D	partially yes	–	weakly yes	–
		A/B vs. C/D	partially yes	no	–	–
	VOT values	English		no	no	yes
		French	–	yes	–	–
		German	–	weakly yes	–	–
		Polish	yes	weakly yes	–	–
III.	MPhA composite score	A vs. B	–	–	yes	yes
		C vs. D	–	no	no	–
		B/C vs. A/D	partially yes	–	no	–
		A/B vs. C/D	–	no	–	yes
	MPhA self-reports		partially yes	yes	yes	–

The following sections will present a summary of the findings of the three series of studies, followed by a discussion thereof in the light of the proposed models of multilingualism.

7.3.7.1. Study I – Summary of the findings

In Study I, the cross-linguistic influence was operationalised as referring to the raters' perception of accentedness and their identification of the subjects' first language on the basis of provided speech samples of the third language. The overall results of the performed foreign accentedness ratings and L1 identification pointed to a prevailing L1 effect and provided some weak but consistent support for the L2 effect.

The data from Study I substantiate the claim that it is the motor routines of the mother tongue that dominate the acquisition of the third language phonology rather than another previously learnt foreign language (cf. Gut 2010; Ringbom 1987). However, there is some evidence of a complementary L2 effect on the acquired L3, based on the L1 identification patterns, although its strength was not that significant.

The prevalence of L1-accentedness of the L3 speech could be attributed, to some extent, to the limitations of this study design, due to which only read rather than spoken speech samples were subjected to the foreign accentedness ratings. The more monitored reading mode in the L3 performance was shown in previous studies to have more phonetic features transferred from the first language, whereas the less monitored speaking mode tended to exhibit more traces of the second language. Consequently, the hypothesised switch to a 'foreign language mode' (cf. Cohen 1995) could have been more evident in the less controlled performance mode than it was the case in the present study.

The role of the typological proximity was not fully confirmed in this study, and the L3 characteristics did not come to the fore as a strong predictor either. Nonetheless, the perceived non-native accentedness in the L3 was demonstrated to be more salient in the case of typologically related pairs of languages rather than more distant pairings, which may be interpreted as an indication of typology acting as a reinforcing factor for the L2 effect to take place. In this respect, the results show some consistency with Rothman's (2011, 2015) Typological Primacy Model (TPM), according to which the selection of the native vs. non-native sources of transfer in the third language acquisition is conditioned by the typological proximity between the target language and the L1 or L2, as perceived by the learner's internal mechanisms. However, the present findings assign a facilitatory rather than determining role to the typological proximity. It stems from the assumptions of the TPM that the unconscious internal mechanisms operate in the L3 acquisition with the aim at transferring knowledge from previously acquired languages, based on optimal selection. An attempt was made to predict the potential transfer mechanisms between the investigated languages (i.e. Polish, French, English and German), on the basis of their phonetic and phonological proximity (see Chapter 4 section 4.7 for a discussion). Nonetheless, the predictions turned out to be quite mixed depending on the selected criteria for typological proximity, which may account for the rather complex pattern of results evidenced in the performed series of L3 accent ratings.

All in all, the findings of Study I partially disconfirm Hammarberg and Hammarberg's (1993, 2005) stance that it is the other non-native language (L2) rather than the mother tongue that constitutes the stronger source language in L3 phonological acquisition leading to an L2-accented speech in L3 performance. Conversely, the generated data is more consistent with the assumption of a combined cross-linguistic influence, which involves simultaneous influence of more than one previously acquired languages on the target language, as proposed by De Angelis (2007). In this way, the results substantiate a necessary extension of the traditional SLA view of a one-to-one transfer between the source and the target language, and an adoption of a wider perspective to account for transfer-related phenomena in multiple language acquisition.

7.3.7.2. Study II – Summary of the findings

As far as Study II is concerned, one of the main findings was that the multilingual participants differentiated between their respective language systems with respect to the VOT durations in the L1, L2 and L3. In addition, the new 'hybrid' categories for the L3 VOT tended to deviate from both the L1 and L2 categories, thus maintaining a phonetic contrast between the three language systems. The reported compromise VOT values in the third language evidenced that the phonetic properties under examination were transferred from L1 Polish and the respective second languages (L2), thus providing clear counterevidence against the first language as the only source of potential cross-linguistic influence in mulltilinguals acquiring a subsequent foreign language and, consequently, contradicting the findings by Ringbom (1987) or Pyun (2005).

The results of the performed cross-linguistic correlations and multiple regression analyses for the L3 VOT values as a dependent variable demonstrated a generally stronger effect of the L2 VOT over the L1 VOT patterns as independent variables, however, the magnitude of this effect differed depending on language groupings. Consequently, this series of studies provided evidence for the co-existence of the L1 and L2 effect on the developing VOT patterns in the third language, thus substantiating the assumption of a combined cross linguistic influence (CLI) in third language acquisition, as suggested by De Angelis (2007). The present findings were found to be largely consistent with previous studies on the VOT patterns in L3 phonological acquisition (Llama et al., 2010; Wunder, 2010, Wrembel, 2010), which pointed to a combined CLI from both the native and non-native languages.

With respect to the role of typology, similar patterns of the VOT performance were found in the typologically related groups as opposed to the more distant groupings. These results were interpreted as a combined effect of the L2 reinforced by the typological closeness between English and German as aspirating languages with respect to the VOT parameter. However, since it is difficult in this case to tease apart the L2 effect from the potential impact of typology, we cannot draw any univocal conclusions as to which factor is the actual predictor of the L3 VOT performance. Further comparisons conducted across the groups with a similar typological distance failed to provide more support for the effect of typology as in some cases the mean VOT durations were found to be very similar irrespective of the typological proximity between the language combinations involved in the language groupings. Summing up, no conclusive evidence of the typological proximity effect can be drawn on the basis of the analysis.

Study II aimed also to investigate if the observed VOT patterns in L3 acquisition parallel those reported in the SLA literature (cf. Flege, 1987; Flege and Eefting, 1988). It was generally found that the intermediate values observed for the L3 VOT differ from the category assimilation reported in the SLA studies, in which the VOT patterns tend to form a hybrid between the native and target values. In the case of the L3 VOT intervals, the hybrid value formation seems to be a compromise between the native L1 VOT, the target VOT, and an intervening variable of the previous foreign language system known to the multilingual participants.

Further, the findings confirmed that the VOT patterns followed the universal effects of the place of articulation (PoA) and the vocalic context, as suggested by Maddieson (1997). The results showed progressively longer VOT values for velars when compared to alveolars and bilabials in all the language systems in all the groups under investigation. The universal vocalic context effects, i.e. longer VOT when a plosive is followed by a high rather than a low vowel, were also generally observed; however, this trend was particularly consistent in the L1 and L2, whereas the L3 proved less stable in this respect. It is noteworthy that by adhering to the language universal effects with regard to the VOT patterns, the external validity of Study II was confirmed.

On the whole, the results of the global analysis of the VOT goodness of fit and the VOT values for particular languages in Study II provided support for combined cross-linguistic influence, with indication of both the L1 effect and partial influence from the L2. There was also evidence of both facilitative and non-facilitative transfer. Further, the typology was reported to have a facilitating effect on the acquisition of the L3 VOT patterns; however, it was not confirmed to be a decisive predictor of success. Moreover, the specific L3 features were found to be strongly influential.

7.3.7.3. Study III – Summary of the findings

In Study III the L3 phonological performance measures included metaphonological awareness composite scores as well as self-reports and metacomments. The study was aimed to develop and test a complex coding system, with a view to analysing quantitatively and qualitatively the data elicited through verbal protocols. Moreover. a new formula for quantifying a composite measure of metaphonological awareness was designed and applied.

The participants showed evidence of various types of metaphonological awareness manifested through the ability to perceive a mismatch between L3 output and input, to notice discrepancies in the auditory form between their L3 production and the target forms, to perform analyses of the phonetic properties of speech in the L3, to formulate phonological rules, etc. The findings supported Schmidt's (2009) noticing hypothesis, according to which conscious awareness is a conditioning factor for learning, since the participants with a higher metaphonological awareness demonstrated also a higher proficiency level in their second and third languages.

The multilingual participants' introspections about perceived influences and interactions between language systems in their repertoires provided strong evidence for the existence of cross-linguistic awareness. The declared instances of the cross-linguistic influence attested both the native and non-native sources of transfer, i.e. L1-to-L3 and L2-to-L3 transfer. The majority of the participants declared their second language (L2) to prevail as the source of CLI in their third language productions, which was in line with the assumption of the foreign language effect, or the L2 status, acknowledged in the literature (cf. Cenoz, 2001; De Angelis, 2007; Hammarberg and Hammarberg, 2005). A number of responses indicating combined cross-linguistic influence offered also considerable support to De Angelis' (2007) proposal of a combined cross-linguistic influence. The explanatory accounts provided by the participants corresponded closely to the factors conditioning CLI identified in the L3 literature such as the chronology and route of acquisition, frequency or recency of language use or typological closeness (cf. Cenoz, 2001; De Angelis, 2007; Hammarberg, 2009). Moreover, the metacognitive comments on multilingual advantage in subsequent language learning reflected the arguments identified by scholars working on the L3 acquisition (cf. Cenoz and Jessner, 2000; Cenoz 2003, De Angelis 2007, Gut 2010).

The effect of the L3 proficiency level on the degree of metaphonological awareness was substantiated in the great majority of the investigated parameters, e.g. correct modifications of the L3 speech, higher rates of reported CLI, greater complexity levels of the comments, and more instances of phonological rules. However, the conducted analyses failed to demonstrate consistent correlations between the parameters of accentedness, comprehensibility, correctness and the composite score of metaphonological awareness in three out of four groups. This indicates that metaphonological awareness, operationalised as the MPhA composite score, does not necessarily contribute to the perception of a smaller degree of foreign accentedness, better comprehensibility and higher pronunciation correctness in the third language. However, such relations were attested in one of the language groupings. Several of the participants' variables were found to be correlated with the MPhA composite score, including the participants' age, the level of proficiency in the L2 and L3, the year of studies, and the number of foreign languages known by the participants.

On the whole, the results pointed to the typological proximity as partially facilitative and to the significance of psychotypology. The results of the two tasks appeared somewhat mixed as far as the role of the first and second languages as sources of CLI in the L3 phonological acquisition is concerned. Nonetheless, there were strong indications of the combined CLI and the L2 status evidenced in the self-reports. Further support was also provided for the impact of the L3 features.

7.3.7.4. Discussion of L3 acquisition models

The findings of the global comparative analysis based on the three conducted series of studies will now be interpreted in the light of the explanatory models designed specifically for multiple language acquisition, which were presented in Chapter 2 (section 2.4). The proposed third language acquisition models include the Cumulative-Enhancement Model (CEM) by Flynn et al. (2004) and Berkes and Flynn (2012); the L2 Status Factor Model by Bardel and Falk (2007, 2012); and the Typological Primacy Model (TPM) by Rothman (2010, 2011, 2013, 2015).

With respect to the general predictions about potential sources of multilingual transfer, including 4 possible scenarios, i.e. (1) no transfer, (2) absolute L1 transfer, (3) absolute L2 transfer, (4) L1 and/or L2 transfer, as stipulated by Rothman (2015: 182), the present findings support the fourth scenario allowing for the interplay of the native and non-native transfer.

As far as the Cumulative-Enhancement Model for Language Acquisition (CEM) as first proposed by Flynn et al. (2004) is concerned, the major research question was whether the first language (L1) of a multilingual speaker maintains its special privileged role in the acquisition of a subsequent foreign language (L3) or if other languages known to a multilingual can also exert an impact on subsequent language acquisition. The present findings evidenced the cumulative effect of the first and the second languages, in accordance with the main tenet of this model, namely that language acquisition is cumulative in nature, thus all the previously known languages can have a potential impact on the subsequent language acquisition. The results question the privileged role of the learners' first language (L1) in the process of multilingual acquisition, as was traditionally attested in second language acquisition.

Another claim maintained in the model referred to only the facilitative nature of transfer, based on the assumed lack of redundancy in linguistic representation and necessarily the enhancement of subsequent language learning. Consequently, the model predicts that any instance of non-facilitative transfer from previous languages would be neutralized or blocked. This claim, however, was not substantiated by the present data since both facilitative and non-facilitative cross-linguistic influence was evidenced. Nonetheless, the motivation for excluding the possibility of non-facilitative transfer has been questioned by some scholars on the grounds that it would pose an unrealistic burden on limited cognitive resources in the process of L3 acquisition.

Recapitulating, the present findings evidenced a combined cross-linguistic influence that was both facilitative and non-facilitative, which can be interpreted as providing partial support for the Cumulative-Enhancement Model or rather for its ‘weak’ version, in the sense that CLI in L3 acquisition is not restricted to either a default L1 or L2.

The framework of the L2 Status Factor Model, as postulated by Bardel and Falk (2007, 2012) or Falk and Bardel (2011), assumed a privileged position of the second language, referred to as the ‘foreign language effect’ or the L2 status factor. This, in turn, indicated that the L2 acts as the prevailing source of transfer, particularly in the initial stages of the acquisition of a third language, independently of any typological similarity between the languages involved.

The hypothesis that the L2 can supersede the L1 as the main source of transfer in L3 acquisition was based on the assumption of a greater cognitive similarity between the L3 and L2, rather than between the L3 and L1. These similarities between the second and third languages are clearly reflected in the participants' profiles data with respect to such factors as the setting and route of acquisition, the age of onset, the years of formal training, metalinguistic knowledge and learning strategies, etc. The inherent differences between the native and non-native languages contribute to the fact that language learners tend to classify them differently and, consequently, co-activate non-native languages in the subsequent language acquisition.

The global findings have failed to provide support for the strong version of the L2 Status Factor Model since, although the second language exerted some impact on various L3 phonological measures in the present series of studies, this influence was mostly cumulative in nature rather than taking precedence over the impact exerted by the first language. A weaker version of this model could put forward the L2 status as a significant predictor of both negative and positive transfer from the second to the third language, based on the neurolinguistic and cognitive similarity between these languages, as opposed to the first language. Such a weaker stance could thus be more realistic in the light of the present findings. Further, a non-facilitative transfer is also predictable, which remains in line with the attested patterns of L3 phonological performance.

What is also appealing in the L2 Status Factor model is the neurolinguistic basis for its predictions stemming from Paradis' psycholinguistic model of bilingualism (2004) and the distinction between the declarative and the procedural memory (Paradis 2004, 2008). As advocated by Paradis (2009), the implicit linguistic competence and explicit metalinguistic knowledge remain neurolinguistically distinct, i.e. they have various memory sources and cerebral representations. The former is sustained by the procedural memory involving non-conscious representations, while the latter relies on the declarative memory with conscious representations. Following Paradis' claim (2008), first language acquisition relies on the implicit mechanism of procedural memory, whereas the later acquired languages (L2, L3, Ln) are sustained, to a large extent, by declarative memory, and therefore, they are more likely to demonstrate dynamic interference from one another rather than from the L1. However, the development of L2 proficiency is accompanied by a gradual shift from the reliance on explicit metalinguistic knowledge towards more implicit competence, which leads to some automatization of the second language performance. The same process may apply to third and additional language acquisition, yet the automatization is bound to be less advanced due to a lesser stability and lower proficiency of the subsequent language system.

In sum, the L2 Status Factor Model seems to be a strong hypothesis that offers too straightforward predictions, whereas the present data does not support this hypothesis and points in the direction of a more complex pattern of acquisition.

The third theoretically-based proposal, namely the Typological Primacy Model (TPM), was put forward by Rothman (2010, 2011, 2013, 2015). Similarly to CEM, the TPM model assumes multiple sources of transfer as well as the access to both the L1 and L2 at the L3 initial state, yet the rationale behind the selection of the sources of transfer differs. Rothman stipulates that this selection depends on the perceived typological proximity of the language pairings involved, and that the typological assessment takes place in the early stages of the third or additional language acquisition.

What is interesting is the motivation behind Rothman's claims, namely the principle of a general cognitive economy, in accordance with which the human mind is predisposed towards the least effort when engaging in a cognitive task. What appears more controversial is the claim that transfer occurs holistically from only one of the previously available language systems (i.e. L1 or L2). The idea of a full transfer from only one language system was not corroborated by the present data. To the contrary, the across-studies analyses demonstrated that a property-by-property or gradual transfer was also feasible. The motivation behind the TPM appears quite appealing as long as the proposed principle of linguistic and cognitive economy is not equated with a complete transfer from one of the available language systems. Although the cognitive burden is obviously considerably more complex in multilingual acquisition and processing, it seems unrealistic that inhibition processes are necessary to suppress completely the activation of the other system, and this claim seems unsubstantiated by the available data.

The present findings provide some evidence in favour of the typologically-based cross-linguistic influence, but this evidence is rather weak and inconsistent. However, no holistic transfer is attested in the data, although the investigation did not concern the initial stages of third language acquisition but was conducted at a more intermediate stage. Unlike the CEM, the TPM allows for both facilitative or non-facilitative transfer, as was evidenced in the present data. Further, Rothman in the TPM attempted to tease apart the factors of typology and L2 status, which was also one of the goals of this dissertation. Nonetheless, the global analyses have failed to separate inconclusively the influence of the typological proximity from that of the L2 status, pointing to more intertwined and structure-dependent relationships. All in all, the series of conducted studies did not corroborate the assumptions of the Typological Primacy Model, demonstrating only some patterns of cross-linguistic influence determined by typological facilitation.

Recapitulating, the existing theoretical models of third language acquisition provide quite conflicting explanatory accounts of this phenomenon. While there seems to be agreement as to the interplay of native and non-native sources of cross-linguistic influence, the predicting factors underlying its strength and directionality remain rather inconclusive. Two of the discussed models consider multilingualism to be conditioned by the cumulative influence of previously acquired language systems, which is corroborated by the present data. However, it seems difficult to provide a unified account for the sources of CLI in the L3 as the process of multilingual acquisition appears to be dynamic and complex. Further, the nature of the attested cross-linguistic influence appears to be gradual and structure-dependent rather than holistic and stable. A question thus appears whether the acquisition of third language phonology deserves a separate explanatory model since the currently existing general L3 acquisition models do not provide satisfactory accounts of the specificity of this process.

Moreover, none of the investigated factors was singled out as predominant in conditioning the cross-linguistic influence in third language acquisition. What becomes evident is that all the language systems of the multilingual user are at play, including the L1, L2 and L3 and that they are conditioned to some extent by the effect of typological relatedness. The process of multilingual acquisition rests on the interplay of these factors. Their mutual impact is gradual and depends on the selected features under study. Different levels of investigations have produced, on the whole, comparative conclusions with some specific indications stemming from each series of studies.

An attempt can be also made to revise some of the existing models of speech in SLA such as Flege's (1995) Speech Learning Model (SLM) or Best's (1995) PAM-L2 and to adapt them to the third language acquisition perspective. The acquisition of a third language does not take place through the filter of the first language alone as there is some evidence of the L2 representations that would need to be allowed for. Following the acquisition of the first foreign language, one may expect that the L1 perceptual space becomes reorganised and reshaped. Yet in the present data the L1 production performance of the multilingual participants (e.g. their VOT durations) did not seem to be that affected by the subsequently learned phonological systems. On the other hand, the L2 values of a rather proficient second language appeared to be quite well established, approximating the respective target native norms. In order to verify fully whether the second language has a facilitative effect on the acquisition of a subsequent system, we would need a control group of second language acquirers, in which the L2 would correspond to the respective L3 in the current language groupings, with the same L1, but without the intervening L2 system. For instance, it would be interesting to compare directly the L3 acquisition to the L2 acquisition with matched language repertoires (e.g. L1 Polish, L2 German, L3 English vs. L1 Polish, L2 English, without the intervening variable of German as the first foreign language). Only such a comparison could provide a fuller answer to the question on the facilitatory effect of the second language on the subsequent phonological acquisition. However, it is almost unrealistic to expect other L2s than English in the Polish educational system, therefore such a design would be rather difficult to implement.

7.3.8. Pedagogical implications

The area of teaching and researching the pronunciation of English as the second language has been widely discussed from the pedagogical perspectives within the SLA framework. A number of recent publications present the state-of-the art overviews of findings as well as achievements and perspectives for the pronunciation pedagogy, featuring edited collections (e.g. Szpyra-Kozłowska et al. 2014, Waniek-Klimczak and Shockey 2013, Waniek-Klimczak and Pawlak 2015, Wrembel et al. 2011), monographs advocating a research-based approach to pronunciation instruction in EFL (cf. Szpyra-Kozłowska 2014) or volumes devoted to the discussion on adequate pronunciation models (cf. Dziubalska-Kołaczyk and Przedlacka 2008). However, relatively little has been published as far as specific guidelines for third language pronunciation pedagogy are concerned, apart from Marx and Mehlhorn (2010) or Wrembel and Sypiańska (2014).

From a more general perspective, Gabryś-Barker (2011, 2013) addressed the specificity of learning a third or subsequent foreign language, claiming that although the process seems similar to learning the previous language, a number of idiosyncractic and contextual factors contribute to the fact that it becomes a qualitatively different experience. She draws particular attention to the complexity of affective functioning of multilingual learners resulting from various dominance areas for particular languages from their multilingual repertoires as well as varying social functions and attitudes. Therefore, Gabryś-Barker's recent publications (2011, 2013) focus on the role of affectivity in formal instructions in multilingual contexts. She proposes that affectivity plays a particularly significant role in adult third or additional language acquisition because adult learners seem more vulnerable in terms of their self-esteem and self-confidence displaying greater levels of anxiety and insecurity than young leaners. Consequently, it is postulated that practical models designed for teaching the L3 to adults should take into account the affective dimensions (Gabryś-Barker 2013: 110-111) by making adult learners aware of the role of their previous learning experiences and allowing for the appraisal thereof and potential modifications of their affectivity.

Along similar lines, Otwinowska and De Angelis (2014), in their edited volume on the teaching and learning in multilingual contexts, present broad sociolinguistic and educational perspectives of this process. The underlying idea is that educators should not fail to take advantage of the prior linguistic knowledge and reactivate it in the multilingual education by enhancing metalinguistic awareness of both the teachers and the learners alike.

Moreover, several pedagogically oriented projects have been launched to promote subsequent foreign language learning and receptive multilingualism featuring, for instance, the DaFnE project (Deutsch als zweite Fremdsprache nach Englisch als erster Fremdsprache / German as L3 after English as L2) (cf. Hufeisen 1994) or EuroComGerm project (Learning to read Germanic languages) (Hufeisen & Marx 2007). Such projects are based on the assumption that language learners should be made aware of the similarities and differences between various languages as well as trained to apply this knowledge in the process of learning an additional foreign language.

In spite of a growing number of publications concerning the general aspects of multilingual education, little has been available to date with respect to the implications for L3 pronunciation pedagogy. One of the major contributions to the field has been Marx and Mehlhorn's (2010) proposal to encourage positive phonological transfer from the second to the third language. The authors recommend, among others, that language instructors should rely on the learners' declarative knowledge (e.g. familiarity with such concepts as vowels, consonants, word stress, intonation, aspiration, transcription, etc.) as well as previously acquired procedural knowledge (i.e. the strategies of learning foreign language pronunciation, the ability to monitor articulatory gestures, etc). The reliance on these two types of linguistic knowledge is expected to encourage cross-linguistic comparisons between phonetic/phonological features of all the languages at the multilingual learners' disposal (L1, L2, L3, Ln) and, thus, to increase metaphonological awareness.

In a brief overview of pedagogical guidelines for third language pronunciation, Wrembel and Sypiańska (2014) conclude that it is of utmost importance to make multilingual learners aware of their special assets featuring broader phonetic repertoires, cross-linguistic consciousness as well as specific language learning strategies, all of which may have a facilitative effect on subsequent phonological learning. More specific proposals for raising phonological awareness, stemming from the conducted research, involve empowering learners with self-monitoring and self-correction strategies aimed at conscious modification of speech production as well as the tools for performing cross-language phonological analysis. A special emphasis should also be placed on developing perceptual abilities and sensitizing learners to foreign sounds in order to aid comprehension in the subsequent language, in pursuit of the idea of receptive multilingualism.

7.3.9. Conclusions and future directions

The aim of the present book has been to bridge the gap in the existing literature and to provide more insights into third language phonological acquisition, a novel and dynamically developing domain, which seems still understudied. Moreover, the book was geared to offer a more comprehensive understanding of the field and to inform the ongoing debates on the complexity of the process of the acquisition of L3 phonology. Specifically, it intended to examine the sources and directionality of cross-linguistic influence, a construct which constituted the focus of the present investigation. The conducted three series of studies were designed to substantiate the role of native and non-native languages in shaping the phonological acquisition patterns in the third language by examining three selected measures of phonetic performance, i.e. the perceived global accent in L3, VOT as a correlate of foreign accentedness and a composite measure of metaphonological awareness. To this end, a holistic approach was pursued combining complementing methodologies of data collection and analysis that included accentedness ratings, acoustic measurements as well as introspective and retrospective oral protocols. With a view to overcoming methodological limitations inherent in L3 studies, the present empirical investigations involved a large subject pool and a mirror design of language repertoires in four different groups of participants.

Furthermore, the contribution was aimed to test the predictions of three selected theoretical models proposed for multilingual acquisition, by applying them to the phonological domain. None of the models under scrutiny was fully corroborated by the generated data. Although weaker versions of the models were capable of accounting for a cumulative impact of previously acquired languages on the L3 phonology, the conditioning factors underlying the strength and source of CLI were found to be rather inconclusive.

A question remains whether the acquisition of the third language phonology deserves a separate explanatory model. It is the intention of the present author to design a proposal of such a model stemming from the empirical evidence produced so far and based on the tenets of the Dynamic Systems Theory (see de Bot 2012 for an overview) or Natural Phonology (e.g. Stampe 1980, Dziubalska-Kołaczyk 2007, 2012), which may be able to offer a greater explanatory potential than the currently discussed models.

In order to account for the observed variability in the multilingual data, one may attempt to assume a more dynamic perspective on language acquisition and processing by embracing the Dynamic Systems Theory. The assumptions of the DST appear very much in line with the present results, which cannot be fully interpreted within the framework of stable models of multilingual acquisition. A dynamically based model would thus be more applicable to explain the observed patterns of gradual cross-linguistic influence that is property dependent and varying degrees of influence of all the language systems at the disposal of the multilinguals. The present findings seem to correspond closely to de Bot's (2012: 91) claims that languages do not act as independent entities but they constitute parts of an integral system. Further, the basic assumptions of this theory support the emergence of complexity through the application of simple mechanisms as well as continuous development of the system over time, the direction of which escapes simple categorical predictions.

Another promising perspective that may provide more explanatory potential to account for the complexity of multilingual acquisition may be that of Natural Phonology (NP). This framework offers considerable advantages over other phonological theories with respect to its applicability to the models of language acquisition (cf. Dziubalska-Kołaczyk 1990). In accordance with its tenets, one may hope to provide insights into the theoretical foundations of third language acquisition as well as the practical implications of this process. The universal nature of the Natural Phonology framework relies on the assumptions that languages are governed by forces that are implicit in human vocalization and perception (i.e. the ease of articulation and the clarity of perception), and that phonological rules have phonetic motivation. Moreover, NP offers a functionalist position embracing the communicative and cognitive orientation of language as well as the conditioning impact of extralinguistic factors. Its claims are based on external evidence stemming from the first and second language acquisition as well as sociophonetics or the study of speech disorders, therefore adding a multilingual perspective may be a welcome extension of this model.

The present contribution has aimed to provide a deeper understanding of this developing area of studies, yet there is still room for further investigations. Future studies should attempt to overcome even more the methodological limitations characteristic for this field due to the complexity of the multilingual acquisition process. The scope of investigations could still be broadened by embracing the production and perception interface, with more research into perceptual aspects of L3 phonological acquisition as well as other research areas including e.g. suprasegmental features. Moreover, the language groupings of multilingual participants could become more diverse and more balanced through the application of the mirror design methodology. Another aspect worth pursuing in the future would be the creation of standard measures of phonological proficiency assessment to ensure greater validity and comparability of related studies. Since the majority of investigations into third language phonology to date has been cross-sectional in nature, further research should involve more longitudinal designs in order to trace the development of L3 phonology from its initial stages. Future research is also needed to corroborate the facilitative effect of bilingualism on the third language acquisition, not just for general proficiency but, specifically, for the area of phonology.

Third language acquisition research seems still in its infancy, yet its significance and ramifications should not be underestimated. It can provide new insights into the process of language learning that go beyond those stemming from investigations into the first or second language acquisition, as postulated by Flynn et al. (2004). Several scholars maintain that investigations into the acquisition of a third language actually contribute towards a more precise understanding of how language is represented in the mind. Finally, it should be reiterated that in the contemporary world, in which multilingualism appears to be the default state of the human condition and linguistic competence, we truly need to be able to account for the process and outcome of multilingual acquisition.

References

Alderson, J. Charles – Clapham Caroline – Steel David. 1997. "Metalinguistic knowledge, language aptitude and language proficiency". *Language Teaching Research* 1/2: 93-121.

Angelowa, T. – Pompino-Marschall Bernd. 1985. "Zur akustischen Struktur initialer Plosiv-Vokal-Silben im Deutschen und Bulgarischen". *Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation der Universität München* 21: 83-96.

Aronin, Larissa – Singleton David. 2008. Multilingualism as a New Linguistic Dispensation. *International Journal of Multilingualism* 5(1): 1-16.

Association for Language Awareness 2012. About. http://www.languageawareness.org/web.ala/web/about/tout.php (date of access: 20 June 2015).

Bardel, Camilla – Falk Ylva. 2007. The role of the L2 in L3 acquisition: the case of Germanic syntax. *Second Language Research* 23(4): 459-484.

Bardel, Camilla – Falk Ylva. 2012. "The L2 status factor and the declarative/procedural distinction". In: Jennifer Cabrelli Amaro – Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Philadelphia/Amsterdam: John Benjamins, 61-78.

Barkley, Sharon. 2010. The acquisition of orthographic-phonological correspondence rules in L2 and L3 Portuguese: Error resolution, interference, and generalizability. Ph.D. dissertation, University of Florida, Gainesville.

Benrabah, Mohamed. 1991. "Learning English segments with two languages". *Actes du XIIème Congrès international des sciences phonétiques*. Aix-en-Provence: Université de Provence. Service des Publications, 334-337

Benson, Phil – Lor Winnie. 1999. "Conceptions of language and language learning". *System* 27: 459-472.

Berkes, Éva – Flynn Suzanne. 2012. "Further evidence in support of the Cumulative-Enhancement Model: CP structure development". In: Jennifer Cabrelli Amaro – Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Philadelphia/Amsterdam: John Benjamins, 143-164.

Best, Catherine. 1995. "A direct realist view of cross-language speech perception". In: Strange, W. (ed.) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 171-204.

Best, Catherine – Tyler Michael. 2007. "Non-native and second language speech perception". In: Bohn Ocke-Schwen – Munro Murray J. (eds.) *Language experience in second language speech learning*. Amsterdam: John Benjamins, 15-34.

Bialystok, Ellen. 1982. "On the relationship between knowing and using linguistic forms". *Applied Linguistics* 3: 181-206.
Bialystok, Ellen. 1994. "Analysis and control in the development of second language proficiency". *Studies in Second Language Acquisition* 16: 157-168.
Bialystok, Ellen. 2001. "Metalinguistic aspects of bilingual processing". *Annual Review of Applied Linguistics* 21: 169-181.
Bialystok, Ellen – Ryan E.B. 1985. "Towards a definition of metalinguistic skill". *Merrill-Palmer Quarterly* 31: 229-251.
Blank, Cintia – Zimmer Márcia. 2009. "A transferência fonético-fonológica L2 (francês) – L3 (inglês): um estudo de caso". *Revista de Estudos da Linguagem* 17(1): 207-233.
Boersma, Paul – Weenick David. 2010. Praat: Doing phonetics by computer, version 5.2.15 http://www.praat.org.
Bono, Mariana. 2011. "Crosslinguistic interaction and metalinguistic awareness in third language acquisition". In: Gessica De Angelis – Jean-Marc Dewaele (eds.) *New Trends in Crosslinguistic Influence and Multilingualism Research*. Bristol: Multilingual Matters, 25-52
Cabrelli Amaro, Jennifer. 2012. "L3 Phonology: An understudied domain". In: Jennifer Cabrelli Amaro – Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Philadelphia/Amsterdam: John Benjamins, 33-60.
Cabrelli Amaro, Jennifer. 2013. *The Phonological Permeability Hypothesis: Measuring regressive L3 influence to test L2 phonological representations*. Unpublished doctoral dissertation, University of Florida.
Cabrelli Amaro, Jennifer. 2014. *Differential L3 influence on L1 and L2 phonological systems*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.
Cabrelli Amaro, Jennifer – Rothman Jason. 2010. "On L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems". *IRAL* 48: 275-296.
Caramazza, A. – Yeni-Komshian G. – Zurif E. – Carbone E. 1973. "The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals". *Journal of the Acoustical Society of America* 54: 421-428.
CEFR. 2011. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Council of Europe, Education and Languages, http://www.coe.int/t/dg4/linguistic/CADRE1_EN.asp (date of access: 9 June 2015).
Cenoz, Jasone. 2000. "Research on multilingual acquisition". In: Jasone Cenoz – Ulrike Jessner (eds.). *English in Europe: The Acquisition of a Third Language*. Clevedon: Multilingual Matters, 39-53.
Cenoz, Jasone. 2001. "The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition". In: Jasone Cenoz – Britta Hufeisen – Ulrike Jessner (eds.) *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Clevedon: Multilingual Matters, 8-20.

Cenoz, Jasone. 2003. "The additive effect of bilingualism on third language acquisition: a review". *International Journal of Bilingualism* 7: 71-88.

Cenoz, Jasone. 2005. "Learning a third language: cross-linguistic influence and its relationship to typology and age". In: Britta Hufeisen – Robert Fouser (eds.) *Introductory L3 readings*. Tübingen: Stauffenberg Verlag, 1-9.

Cenoz, Jasone. 2009. *Towards multilingual education*. Bristol: Multilingual Matters.

Cenoz Jasone – Valencia Jose. 1994. "Additive trilingualism: evidence from the Basque Country". *Applied Psycholinguistics* 15: 197-209.

Cenoz, Jasone – Genesee Fred. 1998. *Beyond bilingualism: Multilingualism and multilingual education*. Clevedon: Multilingual Matters.

Cenoz, Jasone – Jessner Ulrike (eds.). 2000. *English in Europe: the Acquisition of a Third Language*. Clevedon: Multilingual Matters.

Cenoz, Jasone – Hufeisen Britta – Jessner Ulrike (eds.). 2001. *Cross-linguistic Influence in Third Language Acquisition*. Clevedon: Multilingual Matters.

Cenoz, Jasone – Hufeisen Britta – Jessner Ulrike. 2003. "Why investigate the multilingual lexicon?" In: Jasone Cenoz – Britta Hufeisen – Ulrike Jessner (eds.). *The Multilingual Lexicon*, Dordrecht: Kluwer Academic, 1-9.

Chamot, Anna. 1973. "Phonological problems in learning English as a third language". *International Review of Applied Linguistics* 11(3): 243-250.

Clyne, Michael. 2003. *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.

Cohen, Andrew D. 1995. "In which language do/should multilinguals think?" *Language, Culture and Curriculum* 8: 93-98.

Cohen, Andrew D. 1996. "Verbal reports as a source of insights into second language learner strategies". *Applied Language Learning* 7: 5-24.

Cook, Vivian.1992. "Evidence for multicompetence". *Language Learning* 42: 557-591.

Cook, Vivian. 2002. *Portraits of the L2 User*. Clevedon: Multilingual Matters.

Cook, Vivian (ed.). 2003. *Effects of the Second Language on the First*. Clevedon: Multilingual Matters.

Cummins, Jim. 2001. "Instructional conditions for trilingual development". *International Journal of Bilingual Education and Bilingualism* 4(1): 61-75.

Cummins, Jim. 2008. "BICS and CALP*:* Empirical and theoretical status of the distinction". In: Hornberger, Nancy H. *Encyclopedia of Language and Education*. Part 2. Boston, MA: Springer.

Cutler, Ann – Mehler Jacques – Norris Dennis – Segui Juan. 1986. "The syllable's defining role in the segmentation of French and English". *Journal of Memory and Language* 25: 385-400.

Darcy, Isabelle – Mora Joan C. – Daidone Danielle. 2014. "Attention control and inhibition influence phonological development in a second language". *Proceedings of the International Symposium on the Acquisition of Second Language Speech*. *Concordia Working Papers in Applied Linguistics* 5: 115-129.

De Angelis, Gessica. 1999. "Interlanguage transfer and multiple language acquisition: A case study". Paper presented at TESOL 1999, New York City.

De Angelis, Gessica. 2005. “Interlanguage transfer of function words”. *Language Learning* 55: 379-414.
De Angelis, Gessica. 2007. *Third or Additional Language Acquisition*. Clevedon: Multilingual Matters.
De Angelis, Gessica – Selinker L. 2001. “Interlanguage transfer and competing linguistic systems”. In: Jasone Cenoz – Britta Hufeisen – Ulrike Jessner (eds.) *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Clevedon: Multilingual Matters, 42-58.
de Bot, Kees. 1992. “A bilingual production model: Levelt’s speaking model adapted”. *Applied Linguistics* 13(1): 1-24.
de Bot, Kees. 2004. “The multilingual lexicon: Modelling selection and control”. *The International Journal of Multilingualism* 1(1): 17-32.
de Bot, Kees. 2012. “Rethinking multilingual processing: From a static to a dynamic approach”. In: Jennifer Cabrelli Amaro – Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Amsterdam: John Benjamins, 79-93.
de Bot, Kees – Jaensch Carol. 2013. “What is special about L3 processing?” *Bilingualism: Language and Cognititon*, 1-15.
Dewaele, Jean-Marc. 2001. “Activation or inhibition? The interaction of L1, L2 and L3 on the language mode continuum”. In: Jasone Cenoz – Britta Hufeisen – Ulrike Jessner (eds.) *Cross-linguistic Influence in Third Language Acquisition*. Clevedon: Multilingual Matters, 69-89.
Dijkstra, Ton – van Hell Janet G. 2003. “Testing the language mode hypothesis using trilinguals”. *International Journal of Bilingual Education and Bilingualism* 6(1): 2-16.
Doughty, Catherine – Williams Jessica. 1998. “Pedagogical choices in focus on form”. In: Catherine Doughty – Jessica Williams (eds.) *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press, 197-262.
Dziubalska-Kołaczyk, Katarzyna 1990. *A Theory of Second Language Acquisition within the Framework of Natural Phonology.* Poznań: Adam Mickiewicz University Press.
Dziubalska-Kołaczyk, K. 2002. “Conscious competence of performance as a key to teaching English”. In: Ewa Waniek-Klimczak – Patrick J. Melia (eds.). *Accents and Speech in English*. Frankfurt am Main: Peter Lang. 97-106.
Dziubalska-Kołaczyk, Katarzyna. 2007. “Natural Phonology: universal principles for the study of language (insiders meet outsiders)”. In: Jürgen Trouvain – William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, 6-10 August 2007*. Saarbruecken: University of Saarbruecken, 71-75.
Dziubalska-Kołaczyk, Katarzyna. 2012. “Modern Natural Phonology and phonetics”. In: Cyran Eugeniusz – Kardela, Henryk – Szymanek Bogdan (eds.) *Sound, Structure and Sense. Studies in Memory of Edmund Gussmann*. Lublin: Wydawnictwo KUL, 199-210.
Dziubalska-Kołaczyk, Katarzyna – Przedlacka Joanna (eds.). 2008. *English Pronunciation Models: A Changing Scene. A democratic accent for the global village?* Frankfurt: Peter Lang Verlag.

Dziubalska-Kołaczyk, Katarzyna – Wrembel Magdalena (forthcoming). "(Mis)understanding phonological transfer". In: Onysko, Alexander et al. (eds.) *The Polyphony of English Studies: A Festschrift for Allan James*.

Ecke, Peter – Hall Christopher J. 2000. "Lexikalische Fehler in Deutsch als Drittsprache. Translexikalischer Einfluss auf drei Ebenen der mentalen Repräsentation". *Deutsch als Fremdsprache* 1: 31-37.

Ellis, Nick. 1999. "Cognitive approaches to SLA". *Annual Review of* Applied *Linguistics* 19: 22-42.

Ericsson, Anders – Simon Herbert. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

Falk, Ylva – Bardel Camilla. 2011. "Object pronouns in German L3 syntax: Evidence for the L2 status factor". *Second Language Research* 27(1): 59-82.

Fernandes-Boëchat, Marcia Helena. 2007. "The CCR Theory: A cognitive strategy research proposal for individual multilingualism". *Revista Luminária* 8(1): FAFIUV: União da Vitória-PR, http://www.ieps.org.br/luminaria.pdf.

Fernandes-Boëchat, Marcia Helena – Siebeneicher Brito Karim. 2008. "Speaking Models: From Levelt's Monolingual to Williams & Hammarberg's Polyglot". In: Rauber Andreia S. – Watkins Michael A. – Baptista Barbara O. (eds.) *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. Florianópolis, Brazil: Federal University of Santa Catarina, 199-206.

Flege, James E. 1984. "The detection of French accent by American listeners". *Journal of the Acoustical Society of America* 76(3): 692-707.

Flege, James E. 1987. "The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification". *Journal of Phonetics* 15: 47-65.

Flege, James E. 1988. "Factors affecting degree of perceived foreign accent in English sentences". *Journal of the Acoustical Society of America* 84: 70-79.

Flege, James E. 1991. "Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language". *Journal of the Acoustical Society of America* 89: 395-411.

Flege, James E. 1995. "Second-language Speech Learning: Theory, Findings, and Problems". In: Winifred Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language research*. Timonium, MD: York Press, 229-273

Flege, James E. – Eefting Wieke. 1987. "Cross-language switching in stop consonant perception and production by Dutch speakers of English". *Speech Communication*, 6, 185-202.

Flege, James E. – Hillenbrand James. 1987. "Differential use of closure voicing and release burst as cue to stop voicing by native speakers of French and English". *Journal of Phonetics* 15: 203-208.

Flege, James E. – Eefting Wieke. 1988. "Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation". *Journal of the Acoustical Society of America* 83: 729-740.

Flynn, Suzanne – Foley Claire – Vinnitskaya Inna. 2004. "The Cumulative-Enhancement Model for Language Acquisition: Comparing Adults' and Children's patterns of Development in L1, L2 and L3 acquisition of Relative clauses". *The International Journal of Multilingualism* 1(1): 3-16.

Franceschini, Rita. 2009. "The genesis and development of research in multilingualism. Perspectives for future research". In: Larissa Aronin and Britta Hufeisen (eds.), *The exploration of multilingualism*. Amsterdam: John Benjamins Publishing, 27-61.

Fraser, Helen. 2010. "Cognitive Phonology as a tool for teaching pronunciation". In: S. De Knop – F. Boers – T. De Rycker (eds.) *Fostering Language Teaching Efficiency through Cognitive Linguistics*. Berlin: Mouton de Gruyter.

Gabriel, Christoph – Thulke Johanna – Kupisch Tanja. 2014. *Production and perception of non-native accent in French as a foreign language: Multilingual learners with Mandarin Chinese as a heritage language*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.

Gabryś-Barker, Danuta. 2011. "Multilinguals' learning stories: stability and change". In: Ewa Piechurska-Kuciel ─ Liliana Piasecka (eds). 2011. *Variability and Stability in Foreign and Second Language Learning Context.* Cambridge: Cambridge Scholars Publishing, 22-47.

Gabryś-Barker, Danuta. 2012. (ed.) *Cross-linguistic Influences in Multilingual Language Acquisition.* Heldelberg/Berlin, Springer.

Gabryś-Barker, Danuta. 2013. The Affective Dimension in Multilinguals´ Language Learning Experiences, In: Danuta Gabryś-Barker - Joanna Bielska (eds.) *The Affective Dimension in Second Language Acquisition*. Clevedon: Multilingual Matters, 99-111.

Gallardo del Puerto, Francisco – Gómez Lacabex Esther – García Lecumberri Maria Luisa. 2007. "The assessment of foreign accent by native and non-native judges". *PTLC Proceedings*, London, CD-ROM.

Garcia Lecumberri, Maria Luisa – Gallardo del Puerto Francisco. 2003. "English FL sounds in school learners of different ages". In: Maria del Pilar Garcia Mayo – Maria Luisa Garcia Lecumberri (eds.). *Age and the Acquisition of English as a as a Foreign Language. Clevedon: Multilingual Matters*, 115-135.

Gardner, Richard C. – Tremblay Paul F. – Masgoret Anne-Marie. 1997. "Towards a full model of second language learning: An empirical investigation". *Modern Language Journal* 81: 344-362.

Gass, Susan – Selinker Larry (eds.). 1983. *Language transfer in language learning*. Rowley: Newbury House.

Gass, Susan – Mackey Alison. 2000. *Stimulated recall methodology in second language research*. Mahwah: Lawrence Erlbaum Associates.

Gass, Susan – Selinker Larry. 2001. *Second language acquisition: An introductory course*. Hillsdale, NJ: Lawrence Erlbaum.

Gibson, Martha – Hufeisen Britta. 2011. "Perception of preposition errors in semantically correct versus erroneous contexts by multilingual advanced English as a foreign language learners: Measuring metalinguistic awareness". In: Gessica

De Angelis – Jean-Marc Dewaele (eds.) *New Trends in Crosslinguistic Influence and Multilingualism Research*. Bristol: Multilingual Matters, 74-85.

Gombert, Jean E. (1992). *Metalinguistic Development*. New York: Harvester Wheatsheaf.

Grabe, Esther – Low Ee L. 2002. "Durational variability in speech and the rhythm class hypothesis". In: Carlos Gussenhoven – N. Warner (eds.) *Laboratory Phonology 7*. Berlin: Mouton de Gruyter, 515-546.

Green, David. 1986. "Control, activation, and resource: A framework and a model for the control of speech in bilinguals". *Brain and Language* 27: 210-223.

Green, David. 1998. "Mental control of the bilingual lexico-semantic system". *Bilingualism: Language and Cognition* 1: 67-81.

Green, David – Wei Li. 2014. "A control process model of code-switching". *Language, Cognition and Neuroscience* 29(4): 499-511.

Grosjean, François. 1992. "Another view of bilingualism". In: R. I. Harris (ed.) *Cognitive Processing in Bilinguals*. Amsterdam: North Holland, 51-62.

Grosjean, François. 1998. "Studying bilinguals: Methodological and conceptual issues". *Bilingualism: Language and Cognition* 1: 131-149.

Grosjean, François. 2001. "The bilingual's language modes". In: J. Nicol (ed.) *One mind, two languages*: *Bilingual language processing*. Oxford: Blackwell, 1-22.

Gullberg, Marianne (ed.). 2006. "Gestures and second language acquisition. Guest-edited special issue". *International Review of Applied Linguistics* 44(2).

Gurski, Claire. 2006. Voice onset time as a parameter for identification of bilinguals. Proc. 15th Annual Conference of the IAFPA 2006, Göteborg.

Gut, Ulrike. 2009. *Non-native Speech. A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Frankfurt am Main: Peter Lang.

Gut, Ulrike. 2010. "Cross-linguistic influence in L3 phonological acquisition". *International Journal of Multilingualism* 7(1): 19-38.

Gut, Ulrike – Fuchs Robert – Wunder Eva-Maria (eds.). 2015. *Universal or diverse paths to English phonology*. Berlin: Mouton De Gruyter.

Hall, Christopher – Ecke Peter. 2003. "Parasitism as a default mechanism in L3 vocabulary acquisition". In: Cenoz Jasone – Britta Hufeisen – Ulrike Jessner (eds.) *The Multilingual Lexicon*. Dordrecht: Kluwer, 71-86.

Hammarberg, Björn. 1997. "Conditions on transfer in phonology". In: A. James – J. Leather (eds.) *Second Language Speech: Structure and Process*. Berlin: Mouton de Gruyter, 161-180.

Hammarberg, Björn. 2001. "Roles of L1 and L2 in L3 production and acquisition". In: Jasone Cenoz – Britta Hufeisen – Ulrike Jessner (eds.) *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Clevedon: Multilingual Matters, 21-41.

Hammarberg, Björn. 2009. *Processes in third language acquisition*. Edinburgh: Edinburgh University Press.

Hammarberg, Björn – Hammarberg Britta. 1993. "Articulatory re-setting in the acquisition of new languages". *Phonum* 2: 61-67.

Hammarberg, Björn – Hammarberg Britta. 2005. “Re-setting the basis of articulation in the acquisition of new languages: A third-language case study”. In: Britta Hufeisen – Robert Fouser (eds.) *Introductory Readings in L3*. Tübingen: StauFFenburg Verlag, 11-18.

Hansen, Edwards – Jette H. – Zampini Mary L. (eds.). 2008. *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins Publishing Co.

Herdina, Philip – Jessner Ulrike. 1994. The Paradox of Transfer. Paper presented at IRAAL conference, Dublin.

Herdina, Philip – Jessner Ulrike. 2002. *A Dynamic Model of Multilingualism: Changing the Psycholinguistic Perspective*. Clevedon: Multilingual Matters.

Hoffman, Charlotte. 1985. “Language acquisition in two trilingual children”. *Journal of Multilingual and Multicultural Development* 6: 479-495.

Hoffmann, Charlotte. 2001. “Towards a description of trilingual competence”. *International Journal of Bilingualism* 5: 1-17.

Højen, A.D. 2000. “Detection of Danish accent in English by native speakers of Danish”. In: Allan James – Jonathan Leather (eds.) *Proceedings of 4th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2000*. University of Klagenfurt, 80-86.

Honikman, Beatrice. 1964. “Articulatory settings”. In: David Abercrombie – et al. (eds.) *In honour of Daniel Jones*. London: Longman, 73-84.

Hufeisen, Britta. 1994. *Englisch im Unterricht Deutsch als Fremdsprache [English in German class]*. Műnchen: Klett Edition Deutsch.

Hufeisen, Britta. 1998. “Individuelle und subjektive Lernerbeurteilungen von Mehrsprachigkeit. Kurzbericht einer Studie”. *IRAL* 36(2): 121-135.

Hufeisen, Britta. 2000. How do foreign language learners evaluate various aspects of their multilingualism. In: Sigrid Dentler – Britta Hufeisen – Beate Lindemann (eds.) *Tertiär- und Drittsprachen Projekte und empirische Untersuchungen*. Tübingen, Stauffenburg, 223-229.

Hufeisen, Britta – Lindemann Beate. 1997. *Tertiärsprachen. Theorien. Modelle. Methoden*. Tübingen, Stauffenburg.

Hufeisen, Britta – Marx Nicole (eds.). 2007. *EuroComGerm: Die Sieben Siebe. Germanische Sprachen lesen lernen [EuroComGerm: The seven sieves: Learning to read Germanic languages]*. Aachen: Shaker.

Huibregtse, Ineke – de Bot Kees. 2002. “Measuring patterns of recovery, impairment or residual language proficiency in bilingual aphasia”. In: Franco Fabbro (ed.) *Advances in the Neurolinguistics of Bilingualism. Essays in honor of Michel Paradis*. Udine: Forum, 219-237.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

James, Carl 1999. “Language awareness: implications for the language curriculum”. *Language, Culture and Curriculum* 12,1: 94-115.

Jarvis, Scott – Pavlenko Aneta. 2007. *Crosslinguistic Influence in Language and Cognition*. New York/London: Routledge.

Jessner, Ulrike. 1999. "Metalinguistic Awareness in Multilinguals: Cognitive Aspects of Third Language Learning". *Language Awareness* 8: 201-209.

Jessner, Ulrike. 2006. *Linguistic Awareness in Multilinguals. English as a Third Language*. Edinburgh: Edinburgh University Press.

Jilka, Matthias. 2000. The contribution of intonation to the perception of foreign accent. Identifying intonational deviations by means of F0 generation and resynthesis. Unpublished PhD Thesis. Stuttgart University.

Kamiyama, Takeki. 2007. Acquisition of French vowels by Japanese-speaking learners: Close and close-mid rounded vowels. Paper presented at the L3 Phonology Satellite Workshop of ICPhS XVI, Freiburg, Germany.

Keating, Patricia – Mikoś Michael – Ganong III William. 1981. "A cross-language study of range of VOT in the perception of stop consonant voicing". *Journal of the Acoustical Society of America* 70: 1260-1271.

Kellerman, Eric. 1979. "Transfer and non-transfer: where we are now". *Studies in Second Language Acquisition* 2, 37–57.

Kellerman, Eric. 1983. "Now you see it, now you don't". In: Susan Gass – Larry Selinker (eds.) *Language transfer in language learning*. Rowley, MA: Newbury House, 112-134.

Kellerman, Eric. 1984. "The empirical evidence for the influence of the L1 in interlanguage". In: Alan Davies – Clive Criper – Phillip R. Howatt (eds.) *Interlanguage*. Edinburgh: Edinburgh University Press, 98-122.

Kellerman, Eric. 1987. Aspects of transferability in second language acquisition. Unpublished Ph.D. Thesis. University of Nijmegen.

Kennedy, Sara – Trofimovich Pavel. 2010. "Language awareness and second language pronunciation: a classroom study". *Language Awareness* 19 (3): 171-185.

Kivistö-de, Souza, Hanna. 2012. The complex nature of L2 phonological awareness: Methodological issues. Paper presented at the V Seminario Integrado de Pesquisas em Linguıstica, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Klein, Elaine. 1995. "Second versus third language acquisition: Is there a difference?" *Language Learning* 45: 419-465.

Kopečková, Romana. 2013. "Cross-linguistic influence in child L3 instructed phonological acquisition". In: L. Aronin – M. Pawlak (eds.) *Essential Topics in Applied Linguistics and Multilingualism. Studies in Honor of David Singleton*. Second Language Learning and Teaching Series. Heidelberg/New York: Springer.

Kopečková, Romana. 2015. "Differences in the perception of English vowel sounds by child L2 and L3 learners". To appear in: Ulrike Gut – Robert Fuchs – Eva-Maria Wunder (eds.) *Universal or diverse paths to English phonology?* Berlin: Mouton de Gruyter.

Kuhl, Patricia K. – Iverson Paul. 1995. "Linguistic experience and the perceptual magnet effect". In: Strange, Winifred (ed.) *Speech perception and linguistic experience: Issues in cross-language research*. York Press: Baltimore, MD, 121-154.

Kulundary, Valeria – Gabriele Alison. 2012. “Examining the role of L2 syntactic development in L3 acquisition: A look at relative clauses”. In: Jennifer Cabrelli Amaro –Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Philadelphia/Amsterdam: John Benjamins, 195-222.

Kupisch, Tanja – Barton D. – Hailer K. – Kostogryz E. – Lein T. – Stangen I. – Van de Weijer J. 2014. “Foreign Accent in Adult Simultaneous Bilinguals”. *Heritage Language Journal* 11(2): 123-150.

Lambert, Wallace E. – Tucker Richard G. – d’Anglejan A. 1973. “Cognitive and attitudinal consequences of bilingual schooling”. *Journal of Educational Psychology* 65(2): 141-159.

Lasagabaster, David. 1997. “Creatividad y Conciencia Metalinguistica: Incidencia del Aprendizaje del Inglés como L3. PhD dissertation”. Bilbao: Servicio Editorial de la Universidad del País Vasco.

Laver, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lechner, Simone – Kohlberger Martin. 2014. *Phonetic transfer onto L3 English in subtractive bilinguals*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.

Lein, Tatjana – T. Kupisch – Joost van de Weijer. 2015. “VOT and global foreign accent in German-French simultaneous bilinguals”. *International Journal of Bilingualism*, published online.

Leow, Ronald P. 2001. “Attention, awareness, and foreign language behavior”. *Language Learning* 51 (supplement 1): 113-155.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lindemann, Beate. 2000. “Da fa ̈llt mir immer zuerst ein englischer Wort ein.” Zum Einfluß der ersten Fremdsprache beim U ̈bersetzen ins Deutsche”. In: Sigrid Dentler – Britta Hufeisen – Beate Lindemann (eds.) *Tertiär-und Dritssprachen. Projekte und empirische Untersuchungen*. Tübungen: Stauffenburg, 57-65.

Lisker, Leigh – Abramson Arthur. 1964. “A cross-language study of voicing in initial stops”. *Word* 20: 384-422.

Llama, Raquel – Cardoso Walcir – Collins Laura. 2010. “The Influence of Language Distance and Language Status on the Acquisition of L3 Phonology”. *International Journal of Multilingualism* 7(1): 39-57.

Llama, Raquel – López-Morelos Luz Patricia. 2014. *Production of voiceless stops by trilingual children and teens*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.

Llisterri, Joaquim – Poch-Olivé Dolors. 1987. “Phonetic Interference in Bilingual’s Learning of a Third Language”. *Proceedings of the XIth International Congress of Phonetic Sciences*. Tallinn: Academy of Sciences of the Estonian SSR, 134-147.

Long, Michael – Robinson Peter. 1998. "Focus on form: Theory, research and practice". In: C. Doughty – J. Williams (eds.) *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Louriz, Nabila. 2007. "Alignment in L3 phonology". *Langues et Linguistique* 18/19: 129-160.

Lozano, Cristobal. 2002. "The interpretation of overt and null pronouns in non-native Spanish". In: H. Marsden – S. Pourcel – M. Whong-Bharr (eds.) *Durham Working Papers in Linguistics* Vol. 8. Somerville MA: Cascadilla Press, 53-66.

Maddieson, Ian. 1997. "Phonetic universals". In: William J. Hardcastle – John Laver (eds.) *Handbook of Phonetic Sciences*. Oxford: Blackwell, 619-639.

Major, Roy C. 1987. "English voiceless stop production by speakers of Brazilian Portuguese". *Journal of Phonetics* 15: 197-202.

Major, Roy C. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Mahwah, NJ: Erlbaum.

Marian, Viorica – Blumenfeld Henrike K. – Mizrahi Elena – Kania Ursula – Cordes Anne-Kristin. 2013. "Multilingual Stroop performance: Effects of trilingualism and proficiency on inhibitory control". *International Journal of Multilingualism* 10/1: 82-104.

Marx, Nicole – Mehlhorn Grit. 2010. "Pushing the positive: encouraging phonological transfer from L2 to L3". *International Journal of Multilingualism* 7(1): 4-18.

Medina, Almitra D. 2008. Concurrent verbalization, task complexity and working memory: Effects on L2 learning in a computerized task. Ph.D. dissertation, Washington, D.C., Georgetown University.

Meisel, Jürgen. 1983. "Transfer as a second language strategy". *Language and Communication* 3: 11-46.

Missaglia, Federica. 2010. "The acquisition of L3 English vowels by infant German-Italian bilinguals". *International Journal of Multilingualism* 7(1): 58-74.

Mora, Joan Carles – Rochdi Youssef – Kivistö-de Souza Hanna. 2014. "Mimicking accented speech as L2 phonological awareness". *Language Awareness* 23:1-2: 57-75.

Na Ranong, Sirirat – Leung Yan-kit I. 2009. "Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem". In: Leung (ed.) 162-191.

Neuhauser, Sara. 2011. *Foreign accent imitation and variation of VOT and voicing in plosives*. International Congress of Phonetic Sciences XVII, 1462-65.

Odlin, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

Oh, Janet – Jun Sun-Ah – Knightly Leah – Au Terry. 2003. "Holding on to childhood language memory". *Cognition* 86: 53-64.

Oksaar, Els. 1978. "Pre-school Trilingualism: A Case Study". In: Peng, F.C.C. – von Raffler-Engel W. (eds.) *Language Acquisition and Developmental Kinesics*. Hiroshima, 129-138.

Onishi, Hiroshi. 2014. *The effect of L2 experience in L3 perception*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.

Osborne, Andrea. 2003. "Pronunciation strategies of advanced ESOL learners". *IRAL* 41: 131-143.

Otwinowska, Agnieszka – De Angelis Gessica (eds.). 2014. *Teaching and Learning in Multilingual Contexts: Sociolinguistic and Educational Perspectives*. Clevedon: Multilingual Matters.

Paradis, Michel. 1977. "Bilingualism and aphasia". In: H. Whitaker – HA. Whitaker (eds.) *Studies in neurolinguistics* (Vol. 3). New York: Academic Press, 65-121.

Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins Publishing.

Paradis, Michel 2008. "Language and communication disorders in multilinguals". In: Brigitte Stemmer – Harry A. Whitaker (eds.) *Handbook of the Neuroscience of Language*. Amsterdam: Elsevier, 341-349.

Paradis, Michel 2009. *Declarative and procedural determinants of second languages* (Vol. 40). Amsterdam, Netherlands: John Benjamins.

Pearl, E. – Lambert W. 1962. "The relation of bilingualism to intelligence". *Psychological Monographs* 76: 1-23.

Pennington, Martha. 1999. *Equivalence classification in language transfer*. Paper presented at TESOL 1999, New York.

Piske, Thorsten – MacKay I.R.A. – Flege James E. 2001. "Factors affecting degree of foreign accent in an L2. A review". *Journal of Phonetics* 29: 191-215.

Poulisse, Nanda – Bongaerts Theo. 1994. "First language use in second language production". *Applied Linguistics* 15: 36-57.

Pyun, K-S. 2005. "A Model of Interlanguage Analysis – the Case of Swedish by Korean Speakers". In: Britta Hufeisen – Robert Fouser (eds.) *Introductory Readings in L3*. Tübingen: Stauffenberg Verlag, 55-70.

Renou, Janet. 2001. "An examination of the relationship between metalinguistic awareness and second-language proficiency of adult learners of French". *Language Awareness* 10: 248-267.

Riney, Timothy J. – Takagi Naoyuki. 1999. "Global Foreign Accent and Voice Onset Time Among Japanese EFL Speakers". *Language Learning* 49(2): 275-302.

Ringbom, Håkan. 1986. "Crosslinguistic influence and the foreign language learning process". In: Eric Kellerman – Michael Sharwood-Smith (eds.) *Crosslinguistic Influence in Second Language Acquisition*, Oxford: Pergamon Press: 150-162.

Ringbom, Håkan. 1987. *The Role of the Mother Tongue in Foreign Language Learning*. Clevedon: Multilingual Matters.

Ringbom, Håkan. 2002. Levels of transfer from L1 and L2 in L3- acquisition. In: *Proceedings of the Second International Conference on Trilingualism*, I. Ytsma 81 M. Hooghiemstra (eds.). Leeuwarden: Fryske Akademie (CD Rom).

Ringbom, Håkan. 2003. *If you know* Finnish as L2, there will be no major problem learning Swahili. Paper presented at the International Conference on Trilingualism and Third Language Acquisition, Tralee, Ireland, September 4-6.

Rivers, Wilga M. 1979. "Learning a sixth language: an adult learner's daily diary". *The Canadian Modern Language Review* 36: 67-82.

Rivers, Wilga M. (ed.) 1987. *Interactive Language Teaching*. Cambridge: Cambridge University Press.

Roehr, Karen. 2006. "Metalinguistic knowledge in L2 task performance: a verbal protocol analysis". *Language Awareness* 15(3): 180-198.

Roehr, Karen – Gánem-Gutiérrez Gabriela A. 2009. "The status of metalinguistic knowledge in instructed adult L2 learning". *Language Awareness* 18,2: 165-181.

Rothman, Jason. 2010. "On the typological economy of syntactic transfer: Word order and relative clause attachment preference in L3 Brazilian Portuguese". *International Review of Applied Linguistics* 48(2-3): 245-273.

Rothman, Jason. 2011. "L3 Syntactic Transfer Selectivity and Typological Determinacy: The Typological Primacy Model". *Second Language Research* 27(1): 107-127.

Rothman, Jason. 2013. "Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance". In: S. Baauw – F. Dirjkoningen – M. Pinto (eds.) *Romance languages and linguistic theory 2011*, Amsterdam: John Benjamins, 217-247.

Rothman, Jason. 2015. "Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered". *Bilingualism: Language and Cognition* 18(2): 179-190.

Rothman, Jason – Cabrelli Amaro Jennifer. 2010. "What variables condition syntactic transfer? A look at the L3 initial state". *Second Language Research* 26(2): 219-218.

Rothman, Jason – Cabrelli Amaro Jennifer – de Bot Kees. 2012. "Third language acquisition". In: Julia Herschensohn – Martha Young-Scholten (eds.) *The Cambridge Handbook of Second Language Acquisition*, Cambridge: Cambridge University Press, 372-393.

Sampson, Geoffrey – David Gil – Peter Trudgill (eds.). 2009. Language Complexity as an Evolving Variable. Oxford: Oxford University Press.

Sanders, Marianne – Meijers Guust. 1995. "English as L3 in the elementary school". *ITL Review of Applied Linguistics* 107/108: 59-78.

Scheuer, Sylwia. 2000. "What makes foreign accent sound foreign?". In: A. James – J. Leather (eds.) *Proceedings of 4th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2000*. University of Klagenfurt, 306-314.

Schmidt, Richard. 1990. "The Role of Consciousness in Second Language Learning". *Applied Linguistics* 11: 129-158.

Schmidt, Richard. 1993. "Awareness and second language acquisition". *Annual Review of Applied Linguistics* 13: 206-226.

Schmidt, Richard. 2001. "Attention". In: Peter Robinson (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, 3-32.

Selinker, Larry. 1969. "Language transfer". *General Linguistics* 9(2): 67-92.

Shahidi, A.H. – Rahim Aman. 2011. "An acoustical study of English plosives in word initial position produced by Malays". *The Southeaast Asian Journal of English Language Studies* 17(2): 23-33.

Sharwood-Smith, Michael. 1983. "On first language loss in the second language acquirer: problem of transfer". In: Susan Gass – Larry Selinker (eds.) *Language Transfer in Language Learning*. Rowley, MA: Newbury House, 222-231.

Sharwood-Smith, Michael. 1994. *Second Language Acquisition: Theoretical Foundations*. London: Longman.

Sicola, Laura. 2010. "Attention to phonological form; Modified output in task-based interaction". In: Martin Pütz – Laura Sicola (eds.) *Cognitive processing in Second Language Acquisition*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 334-350.

Siegel, Jeff. 2003. "Social context". In: Catherine J. Doughty – Michael H. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell Publishing, 178-223.

Singh, R. – Carroll S. 1979. "L1, L2 and L3". *Indian Journal of Applied Linguistics* 5: 51-63.

Slabakova, Roumyana 2015. *The Scalpel Model of Third Language Acquisition*. Paper presented at the Workshop on Multilingual Language Acquisition, Processing and Use, 20-21 March 2015, Tromsø, Norway.

Sopata, Aldona. 2013. "Dwujęzyczność, trójjęzyczność, wielojęzyczność: podobieństwa i różnice" [Bilingualism, trilingualism, multilingualism: similarities and differences]. In: Stanisław Puppel and Teresa Tomaszkiewicz (eds), *Scripta manent – res novae*. Poznań. Wydawnictwo Naukowe UAM. 399-408.

Stampe, David. 1980. *A Dissertation on Natural Phonology*. New York: Garland.

Sypiańska, Jolanta. 2013. *Quantity and quality of language use and attrition of L1 Polish due to L2 Danish and L3 English*. Unpublished PhD thesis, AMU Poznań.

Sypiańska, Jolanta. 2014. *Multilingual acquisition of vowels in L1 Polish, L2 Danish and L3 English*. Paper presented at SLE 2014 Workshop on Advances in the Investigation of L3 Phonological Acquisition.

Szpyra-Kozłowska, Jolanta. 2014. *Pronunciaiton in EFL instruction, A Research-Based Approach*. Bristol: Multilingual Matters.

Szpyra-Kozłowska Jolanta, Radomski Marek. 2012. "The perception of foreign-accented Polish – a pilot study". *Research in Language* 10(1), 97−110.

Szpyra-Kozłowska, Jolanta – Guz Ewa – Steinbrich Piotr – Święciński Radosław (eds.). 2014. *Recent developments in applied phonetics*. Lublin: Wydawnictwo KUL.

Tremblay, Marie-Claude. 2007. *L2 influence on L3 pronunciation: Native-like VOT in the L3 Japanese of English-French bilinguals*. Paper presented at the Satellite Workshop of ICPhS XVI, Freiburg, Germany, 3-4 August 2007.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Venkatagiri, H.S. – Levis John M. 2007. "Phonological awareness and speech comprehensibility: an exploratory study". *Language Awareness* 16(4): 263-277.

Vildomec, Věroboj. 1963. *Multilingualism*. Leyden: Sythoff.

Vingerhoets, G. – Van Borsel J. – Tesink C. – et al. 2003. "Multilingualism: an fMRI study". *Neuroimage* 20: 2181-2196.

Waniek-Klimczak, Ewa. 2011. "Aspiration and style: A sociophonetic study of the VOT in Polish learners of English". In: Wrembel Magdalena – Kul Małgorzata –

Dziubalska-Kolaczyk Katarzyna (eds.) *Achievements and perspectives in SLA of speech*. Frankfurt: Peter Lang, 303-316.

Waniek-Klimczak, Ewa – Shockey Linda (eds.). 2013. *Teaching and researching English accents in native and non-native speakers*. Berlin Heidelberg: Springer Verlag.

Waniek-Klimczak, Ewa – Pawlak Mirosław (eds.). 2015. *Teaching and researching the pronunciation of English*. Berlin Heidelberg : Springer-Verlag.

White Paper on Education and Training: Teaching and Learning: Towards the Learning Society, Brussels: EU Commission, COM (95) 590 final, 29 November 1995.

White, Joanna – Ranta Leila. 2002. "Examining the interface between metalinguistic task performance and oral production in second language classrooms". *Language Awareness* 11(4): 259-290.

Williams, Sarah – Hammarberg Björn. 1998. "Language switches in L3 production: implications for a polyglot speaking model". *Applied Linguistics* 19: 295-333.

Wissing, Daan. 2005. "Aspiration of English voiceless stop consonants in Southern Sotho: a case study". *South African Journal of African Languages* 25(3): 189-2005.

Wrembel, Magdalena. 2005. *Phonological metacompetence in the acquisition of second language phonetics*. Unpublished Ph.D. thesis. Adam Mickiewicz University, Poznan, Poland.

Wrembel, Magdalena. 2009. "The impact of voice quality resetting on the perception of a foreign accent in Third Language Acquisition". In: Rauber Andreia S. – Watkins Michael A. – Baptista Barbara O. (eds.). *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Cambridge Scholars Publishing, 291-307.

Wrembel, Magdalena. 2010. "L2-accented speech in L3 production". *International Journal of Multilingualism*. Special Issue 7(1): 75-90.

Wrembel, Magdalena. 2011a. "Cross-linguistic influence in third language acquisition of voice onset time". In: Wai-Sum L. – Zee E. (eds.) *Proceedings of the 17th International Congress of Phonetic Sciences ICPhS*. 17-21 August 2011. Hong Kong. *CD-ROM*. Hong Kong: City University of Hong Kong, 2157-2160.

Wrembel, Magdalena. 2011b. "Metaphonetic awareness in the production of speech". In: Pawlak Mirosław – Waniek-Klimczak Ewa – Majer Jan (eds.). *Speaking and Instructed Foreign Language Acquisition*. Clevedon. Multilingual Matters, 169-182.

Wrembel, Magdalena. 2012a. "Foreign accentedness in Third Language Acquisition; the case of L3 English". In: Jennifer Cabrelli Amaro – Suzanne Flynn – Jason Rothman (eds.) *Third Language Acquisition in Adulthood*. Amsterdam: John Benjamins, 281-309.

Wrembel, Magdalena. 2012b. "Foreign accent ratings in third language acquisition; the case of L3 French". In: Ewa Waniek-Klimczak – Linda Shockey (eds.) *Teaching and researching English accents in native and non- native speakers*. Berlin Heidelberg: Springer Verlag, 29-45.

Wrembel, Magdalena. 2013. “Metalinguistic awareness in third language phonological acquisition”. In: Karen Roehr – Gabriela A. Gánem-Gutiérrez (eds.) *The Metalinguistic Dimension in Instructed Second Language Learning*. London: Bloomsbury, 119-143.

Wrembel, Magdalena. 2014. “VOT Patterns in the Acquisition of Third Language Phonology”. *Concordia Papers in Applied Linguistics (COPAL)* 5: 751-771.

Wrembel, Magdalena. 2015a. “Metaphonological awareness in multilinguals; a case of L3 Polish”. *Language Awareness* 24(1): 60-83.

Wrembel, Magdalena. 2015b. “Cross-linguistic influence in second vs. third language acquisition of phonology”. In: Ulrike Gut – Robert Fuchs – Eva-Maria Wunder (eds.) *Universal or diverse paths to English phonology*. Berlin: Mouton De Gruyter, 41-70.

Wrembel, Magdalena – Gut Ulrike – Mehlhorn Grit (eds.). 2010. “Special Issue on L3 Phonology”. *International Journal of Multilingualism* 7(1): 1-90.

Wrembel, Magdalena – Kul Małgorzata – Dziubalska-Kołaczyk Katarzyna (eds.). 2011. *Achievements and Perspectives in SLA of Speech: New Sounds 2010*. Volume I & II. Frankfurt am Main: Peter Lang.

Wrembel, Magdalena – Sypiańska Jolanta. 2014. “New perspectives on acquiring foreign language pronunciation”. In: Jolanta Szpyra-Kozłowska – Piotr Steinbrich – Ewa Guz – Radosław Święciński (eds.) *Recent Developments in Applied Phonetics*, Lublin: John Paul II Catholic University of Lublin Press, 239-255.

Wunder, Eva-Maria. 2010. “Phonological cross-linguistic influence in third or additional language acquisition”. *Proceedings of 6th New Sounds 2010*, Poznań, 566-571.

Wunder, Eva-Maria. 2011. “Cross-linguistic Influence in Multilingual Language Acquisition: Phonology in Third or Additional Language Acquisition”. In: Gessica de Angelis – Jean-Marc Dewaele (eds.) *New Trends in Crosslinguistic Influence and Multilingualism Research*. Clevendon: Multilingual Matters, 105-128.

W poszukiwaniu nowej perspektywy: międzyjęzykowe wpływy w nabywaniu fonologii trzeciego języka

STRESZCZENIE

We współczesnym świecie wielojęzyczność jest normą raczej niż wyjątkiem, z tego względu badania nad nabywaniem kolejnych języków obcych nabierają obecnie szczególnego znaczenia i wydają się nieodzowne. Książka podejmuje temat akwizycji fonologii w języku trzecim, co wpisuje się w nowatorską perspektywę badawczą dokonującą rozróżnienia między uczeniem się pierwszego języka obcego, a nabywaniem kolejnych. Ponadto przedstawiona została złożoność tego procesu i warunkujących go czynników, m.in. typologii i psychotypologii, poziomu znajomości języków źródłowych i docelowych, różnic jakościowych i ilościowych dotyczących użycia języka czy instruktarzu. Przegląd dotychczasowych badań w tej młodej dziedzinie, obejmujących dwa ostatnie dziesięciolecia, dostarcza więcej pytań niż odpowiedzi.

Głównym przedmiotem badań w obecnej monografii jest specyfika wpływów międzyjęzykowych (ang. *cross-linguistic influence*) w akwizycji fonologicznej w języku trzecim. Zastosowany koncept jest szerszy od powszechnie wcześniej stosowanego terminu transferu i uwzględnia możliwość istnienia wielu źródeł wpływów jak i różnego ich kierunku. Celem badawczym jest weryfikacja założeń oraz hipotez wynikających ze współczesnych modeli teoretycznych zaproponowanych z perspektywy wielojęzyczności. Wybrane modele nabywania języka trzeciego obejmują (1) model promujący priorytetową rolę języka drugiego nad językiem ojczystym (*L2 Status Model* autorstwa Bardel i Falk), (2) model propagujący tzw. wzmocnienie kumulacyjne wszystkich języków nabywanych uprzednio (*Cumulative Enhancement Model* autorstwa Flynn et al.) oraz (3) model prymatu typologicznego, według którego bliskość strukturalna między językami determinuje wybór źródła wpływów (*Typological Primacy Model* autorstwa Rothman).

W powyższym celu przeprowadzone zostały trzy serie badań własnych na czterech odpowiednio dobranych grupach uczestników ze zróżnicowanym repertuarem języków drugich i trzecich z zastosowaniem tzw. paradygmatu odbicia lustrzanego (L2 vs. L3). Badanie pierwsze obejmuje ocenę obcego akcentu, zrozumiałości i poprawności wymowy w języku trzecim oraz identyfikacji języka ojczystego w próbkach mowy w L3 dokonywaną przez grupy oceniających (ang. *raters*). Badanie drugie koncentruje się na

pomiarach akustycznych parametru VOT we wszystkich językach z repertuaru uczestników (L1, L2 i L3) oraz porównaniu uzyskanych wartości między językami jak i z wartościami referencyjnymi dla danych języków. Badanie trzecie dotyczy stopnia świadomość metafonologicznej uczestników określanego na podstawie wypracowanej formuły kalkulacyjnej uwzględniającej aspekty jakościowe i ilościowe danych wygenerowanych poprzez zastosowane protokoły ustne.

Wyniki zostały zanalizowane osobno dla każdego z serii badań jak i globalnie z uwzględnieniem analiz porównawczych między grupami oraz poszczególnymi badaniami. Rezultaty przeprowadzonych badań wskazują na podwójne źródło wpływów międzyjęzykowych w zakresie fonologii języka trzeciego obejmujących zarówno język ojczysty jak i pierwszy język obcy, warunkowane częściowo przez bliskość typologiczną poszczególnych zestawów języków. Ponadto charakter tego zjawiska jest zależny od badanego aspektu, jest gradalny i na tyle złożony, że dotychczasowo zaproponowane modele akwizycji języka trzeciego mogą tylko częściowo wyjaśnić jego specyfikę.

WYDAWNICTWO NAUKOWE UNIWERSYTETU IM. ADAMA MICKIEWICZA W POZNANIU
61-701 POZNAŃ, UL. FREDRY 10
www.press.amu.edu.pl
Sekretariat: tel. 61 829 46 46, fax 61 829 46 47, e-mail: wydnauk@amu.edu.pl
Dział sprzedaży: tel. 61 829 46 40, e-mail: press@amu.edu.pl

Wydanie I. Ark. wyd. 20,5. Ark. druk. 27,0

DRUK I OPRAWA: UNIDRUK, LUBOŃ, UL. PRZEMYSŁOWA 13