Optimizing Cognitive Processing in the L2 Classroom: New Integrated Grammar Teaching Framework, Tested on Unprepared Oral Production of Russian Case Inflection

Ms. Natalia Vladimirovna Parker

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of Leeds

School of Language, Cultures and Societies

31 October 2022

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

ACKNOWLEDGEMENTS

First of all, I would like to say the biggest thank you to my first supervisor Dr. James Wilson for supporting this project from the very start (in fact even before it officially started), for encouraging me to pursue my research and for his titanic efforts to help me secure my funding. Most importantly, this project would not have been possible without the Competitive Studentship awarded to me by Arts and Humanities Research Council (AHRC), which has made an enormous difference. The research training and the tireless support, provided by White Rose College of Art and Humanities (WRoCAH) has made my journey through my PhD exceptionally productive and most enjoyable. Thank you, WRoCAH, for making me feel part of the team. Next thank you goes to my second supervisor Dr. Clare Wright for her invaluable advice and priceless suggestions which led me through the rocky road of SLA. I feel extremely fortunate with both of my supervisors and have to thank them for their inexhaustible patience in answering my endless queries and trying to understand my way of thinking. I am also very grateful to the staff at Language, Cultures and Societies, who organised and led the doctoral training for our cohort and was otherwise involved in this project.

In addition, I would like to thank our post-graduate office for helping me with my paperwork, as well as for their inventiveness in trying to process claims with supporting documents written in Cyrillic. For obvious reasons, I have to thank my participants for volunteering for this project, for their interest in Russian and for their dedication in learning with me. Finally, there are no words which can express my gratitude to my daughter Alexandra Parker for her understanding, help and interest in what I do. Moreover, special thanks goes to her for drawing initial sketches for both Comics tests. Finally, a big Russian thank you goes to my friends in neighbours who helped me manage my life, while most of the last four years was dedicated to this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1
INTRODUCTION

How could teachers teach so that learners learn… more effectively?
(allusion to Dick Allwright, 1984)

In the present global context, learning languages structurally different from English is very topical (British Council Report, 2014; MLA Report, 2007). However, identifying the most effective ways of learning and teaching languages is still a subject of an on-going debate, which has now stopped being the sole realm of language pedagogy and involves research in other disciplines, such as Second Language Acquisition (SLA) and psycholinguistics. In the present study, this issue is analysed from a different perspective – that of leaner's processing, as a lot of recent research in second language (L2) learning and acquisition appears to point in this direction (e.g., Juffs & Harrington, 2011; Kormos & Sáfár, 2006; Lightbown, 2008; Mackey et al., 2002; Martin and N.Ellis; 2012; VanPatten, 2004).

In order to do that, I identify two models that would be suitable for my analysis of the processing mechanisms involved in L2 learning. As processing of information is explained by cognitive psychology, I plan to draw on some developments in this discipline, particularly on Working Memory research, because it is often linked to the success of learning, including that of L2 (e.g., Kormos, 2014; Pawlak, & Biedroń, 2021; Pili-Moss et al., 2020). Furthermore, two influential SLA hypotheses, which are widely recognised as facilitating learner's acquisition in the L2 classroom, namely, Noticing Hypothesis (Schmidt, 1990) and Interaction Hypothesis (Long, 1996), will also be examined from the perspective of the effectiveness of learners' processing.

From the numerous aspects of language that need to be acquired by an L2 learner, I have chosen to investigate grammar for a number of reasons. First, the mapping between meaning and language form has been identified as one of the major problems in L2 acquisition by SLA (e.g., Bates & MacWhinney, 1989; N. Ellis, 2005; Lightbown & Spada, 2006; McManus, 2019; Slabakova, 2006; VanPatten et al., 2004). Second, grammar has been at the centre of a traditional L2 syllabus within language education (e.g., Ur, 2011; Van Patten, 2003), thus, if I am to propose a possible way of making L2 teaching more effective, this appears to be a principal aspect to address. Third, the Focus-on-meaning and Focus-on-form approaches questioned the dominant position of grammar in teaching, which was termed as Focus-on-FormS, (e.g., Doughty & Williams, 1998; R. Ellis et al., 2002; Lightbown

& Spada, 1990; Long, 1991; Loewen, 2011; Richards, 2005). That has resulted in a range of studies investigating their application within L2 instruction, more recent of which suggest the facilitative effect of their integration; this giving one more dimension to the present study. As a continuation of this line of research, there has developed a discussion of the role of explicit and implicit in teaching and learning (e.g., DeKeyser, 2003; N. Ellis, 2005; R. Ellis, 2008; Han & Finneran, 2014; Hulstijn, 2005), which provides some valuable insights into the effectiveness of these two processes inter-connected within an L2 classroom; thus, need to be used to the most benefit of L2 learners.

In addition, this body of research is connected to the concepts of declarative and procedural knowledge, which are operated by cognitive psychology for the explanation of the acquisition of skills. As the main purpose of L2 learning, from a learners point of view, is acquiring language skills, I look at how this process is viewed by the Skill Acquisition Theory. Of the four language skills, learners often consider speaking to be the most difficult, as well as the most desirable to acquire (European Commission survey, 2012). This is also confirmed by Kormos (2014), who said that “when learning a second language, one of the most frequent aims is being able to speak the language, and the acquisition of other skills as reading and writing is seen to be secondary to speaking.” (Kormos, 2014: xvii). That is why, the current study concentrates on the acquisition of grammar in speech, that is, in oral production. Most importantly, I compare the processing suggested to take place during the production of speech (e.g., Bock & Levelt, 2002) with that involved in performing other skills (e.g., Anderson, 2015; DeKeyser, 2015) to identify the differences, which I believe are essential for the effective production of grammar features in speech. These also could possibly show why the Skill Acquisition Theory might not fully explain L2 acquisition (see discussion in 3.2).

To narrow down the scope of my investigations, I decided to focus on the acquisition of inflection, which is part of functional morphology, primarily because it is identified as the most difficult part of L2 grammar to acquire (e.g., DeKeyser, 2005; Slabakova, 2006), while at the same time “being at the core of language acquisition” (Slabakova, 2014: 7). Also, inflection appears to incorporate the two elements which have been opposed in L2 learning and teaching research, that is, meaning and form, which presents an interesting potential for me to see how these are handled by learners’ processing mechanisms. Then, it has been suggested that nominal inflection, for example, “has so far not been investigated at the same level of detail” as “clause structure and the morphosyntax of verbs” (Parodi, Schwartz, & Clahsen, 2004). This is particularly relevant for languages like Russian, where nominal inflection is very rich and complex to acquire. As well as briefly touching on how inflection acquisition is seen by two main SLA approaches, namely, generativists and functionalists, I

analyse the processing aspects specific to inflection, in order to identify how these could impact the effectiveness of its production, which is largely unresearched area.

Despite the abundance of research on language learning, a number of researchers point out that not much of it has been implemented into actual language teaching practices (see Larsen-Freeman, 2015, for discussion; also, Ellis & Shintani, 2013; Johnstone, 2004). One of the explanations for this phenomenon was offered by Ellis and Shintani (2013), who observed that "teachers' starting point, understandably, is not 'How do learners learn?' but rather 'How should I teach?'" (Ellis & Shintani, 2013: 321). In order to address this issue, they suggested "considering teaching in the light of how learners learn" (Ellis and Shintani, 2013: 321). Following this suggestion, in the present study, I will examine how L2 teaching could address and incorporate findings of research on L2 learning, as well as L2 acquisition and learners' processing, and propose a possible way of making L2 teaching more effective by aligning it with these findings. Furthermore, it has been noted that the studies which have been conducted on language learning are mainly "small-scale classroom experiments designed for SLA research purposes. They do not generally appear to have been incorporated into the cluster of pedagogical principles and practices" (Johnstone, 2004: 667). Therefore, I aim to fill this niche by designing a new teaching framework based on distinct principles drawn from the relevant research.

As the basis for my new innovative teaching framework, I have selected Spiral Curriculum which was proposed by Jerome Bruner, often considered to be one of the founders of cognitive psychology, and which gave rise to a number of successful pedagogical innovations, including those in language teaching. To enhance the robustness of the spiralling framework, I developed it further using recent findings in psychology, psycholinguisics and SLA, explaining its principles from the angle of learners' processing. It is designed to provide the conditions necessary for optimizing learners' processing of L2 grammatical features and allows to restructure the teaching instruction to suit learners' restricted cognitive resources, enabling them to use these in the most effective way. I will show how my teaching framework addresses the inferences that I have drawn from the research, discussed above. More specifically, I will explain how my proposed framework could be implemented within the constraints of a standard beginners' language course, using Russian case inflection as an example. Russian nominal case system, notorious for its complexity, including fusion of different grammatical categories and syncretism (when one inflection is recycled for different cases within one paradigm, for example, nominal), presents a perfect ground for testing the effectiveness of my proposed teaching framework. Not surprisingly, L2 learners' difficulties in acquiring Russian case are well-attested (e.g., Arnett

& Lysinger, 2013; Gor, Chrabaszcz & Cook, 2019; Izurin, 2013; Janda, 2002; Rifkin, 2005). Moreover, despite a lot of research interest in case inflection, "processing of nominal inflection […] has remained largely unexplored" (Gor, Chrabaszcz & Cook, 2019: 7). In order to identify how learners' processing is affected by the specifics of Russian case functions and case forms, in the present study, I carried out some analysis of Russian case in view of the two models that I identified in my research literature review for approaching L2 learners' processing. Furthermore, I briefly examined how these processing issues are attended to in Russian beginner textbooks, with idea of building up, in my proposed teaching framework, on what has been already done.

The main aim of my empirical investigations is to test the effectiveness of my proposed teaching framework with regard to beginner learners' production of Russian nominal case inflection in a language classroom. In order to measure learners' performance on case inflection, I had to address the problem of testing grammar in speech (as grammar is normally tested in written tests) by designing two new testing tools, which are described in my Research Design, and which have proved to be invaluable for collecting my data. Within my investigations, I also looked at identifying some of the factors that affect learners' processing in the classroom, to see whether these could shed more light on how grammar teaching could be adjusted to facilitate learners' acquisition of inflection in speech within the classroom environment.

The structure of this thesis is to reflect the topics discussed above as follows. The literature review is divided into two chapters, which address two different directions, namely, research in L2 teaching and learning (Chapter 2); and research relevant to processing in a classroom (Chapter 3). In Chapter 4, I outline my proposed teaching framework, preceding it with a brief excursus of Russian case inflection. My Research Questions for the empirical part of the thesis and my research design, as well as research methodology and data collection procedures, are described in Chapter 5. Finally, Chapter 6 provides the results of my investigations, while Chapter 7 presents their discussion in the light of the literature reviewed, before the conclusions are drawn in Chapter 8.

The present study aims to identify a possible way of increasing the effectiveness of L2 grammar teaching by bringing together research in language learning and learners' processing, as well as skill acquisition studies, and putting these in context of language beginner's classroom by proposing and testing an innovative grammar teaching framework. This way, my research has the potential to bridge the gap between linguistics research and language teaching practices. In the long run, it also could improve the standard of learning Russian as a second language and enhancing the student experience within the HE

curriculum. Furthermore, the study contributes to the SLA research in inflection acquisition and could possibly inform the teaching of other morphologically complex languages. Overall, it attempts to offer an answer to the question posited at the start of this thesis - How could teachers teach so that learners learn… more effectively?

Chapter 2

L2 GRAMMAR TEACHING, LEARNING AND ACQUISITION

As the present study is dedicated to investigating more effective ways of teaching inflection to L2 learners, it appears logical to, first, examine some existing approaches to the teaching of grammar in classroom settings, as well as issues around the acquisition of morphosyntax by L2 learners. In the terms of grammar and syntax are used somewhat interchangeable. Therefore, I will start with brief reviews of some relevant findings in the three fields, namely, language pedagogy, SLA and psycholinguistics, outlining how they inform the proposed grammar teaching framework, before focusing specifically on those processes that are involved in the acquisition of inflection.

2.1. The teaching of L2 grammar: form-meaning relationship

The formal teaching of grammar, where form was paramount, after centuries of dominating language pedagogy, has been nudged out from the research agenda by Communicative Language Teaching (CLT) that focuses on meaning rather than form, often producing learners with exceptional fluency but with inadequately low linguistic competence (e.g., R. Ellis, 2006; Harley & Swain, 1984; Lightbown & Spada, 1990; among others). At the same time, grammar has typically remained central within traditional modular syllabi (e.g., Larsen-Freeman, 2015; Ryzhova, 2008; Ur, 2011; Van Patten, 2003). Usually, grammar tends to be studied separately from speaking skills and learners often struggle to "internalize" (R. Ellis 2006: 84) grammar knowledge, that is, automatically produce various inflections in their speech (see meta-analysis by Norris & Ortega, 2000; also Ryzhova, 2008). Thus, "the tension between the desirability of communicative use of the foreign language, on the one hand, and the felt need for a linguistic focus in language learning, on the other" (Long, 1991: 41) still has not been resolved (see discussion in Mochizuki & Ortega, 2008).

The opposition of "form-focused" instruction (when learners' attention is drawn to language form, as in Focus-on-FormS) and "meaning-focused" pedagogical models (when grammatical form is not part of the instruction) (see Fotos, 2001; and Loewen & Sato, 2018; for an overview) triggered the distinction between explicit and implicit teaching. The former occurs when the rules are explained directly, while the latter aims to get learners to produce required forms without verbalizing the actual rules (R.Ellis, 2005). The first approach can be clearly exemplified by the PPP model, where the three Ps stand for "presentation", "practice", and "production". Originally proposed by Byrne (1986), this model suggests that each grammar feature is to be explicitly explained to learners, than practiced in grammar

exercises (e.g., filling the gaps or opening the brackets); and finally produced orally. The second approach can be illustrated by the Task-Based Instruction (TBI), initially introduced by Prabhu (1987), often referred to as task-based language teaching (TBLT) (R.Ellis, 2017). TBI activities focus on communication, engaging learners in tasks that encourage them to interact; learners are expected to "work out" the rules by using the language with no help from the teacher.

After the initial standoff period between the two approaches above, when the substantial body of research produced mixed results (see review by Ur, 2011), providing critiques, as well as support for either point of view, a slow but distinctive move towards integration has been emerging, with proponents of each approach attempting to assimilate elements from their opponents' models. The problem of integrating the two approaches has even been pronounced as a "central dilemma" of language pedagogy (Richards, 2002).

On the one hand, Long (1991) suggested Focus-on-Form as a feature within CLT, allowing "brief interruptions" in communication to explicitly discuss grammatical aspects encountered (Long, 1991: 46); that was also supported by Swain (1995) and Doughty (2001). The concept was systematised by R.Ellis et al (2002), who defined focus-on-form as "the treatment of linguistic form in the context of performing a communicative task" (R.Ellis et al., 2002; 419). They also emphasized the distinction between "focus-on-form" and "focus-on-forms", with the latter referring more or less to the systematic teaching of grammatical forms within a traditional linguistics-based syllabus. Moreover, in their paper, R.Ellis et al (2002) provide several classifications of the focus-on-form methodologies – explicit and implicit, planned and incidental, conversational and didactic, but most are grouped into "reactive", when a teacher provides feedback (that is, "reacts") for a grammatical error produced, and "pre-emptive", when a particular linguistic form is chosen *in advance* for learners to focus on during their communicative activities. Mochizuki and Ortega (2008) implemented the latter type of focus-on-form in their empirical study (explicitly entitled "Balancing communication and grammar in beginning-level foreign language classrooms") in the form of "guided planning", that is, explicit guidance for a particular grammatical feature (English relative clauses) during the pre-task planning in preparation for a communicative task. A comprehensive review of the developments of focus-on-form can be found in Loewen (2011).

On the other hand, the proponents of form-focused approach started adapting traditional explicit grammar teaching to accommodate CLT needs. For example, Terrell (1991) introduced the term of Explicit Grammar Instruction (EGI) and discussed its role in a communicative approach as "an aid to the learner in the acquisition process" (Terrell, 1991:

62). Then, the concept of “pedagogical grammar” was put forward, encouraging teachers to select grammar aspects necessary for students’ communication (Odlin, 1994), with the idea applied to practical teaching by Savage et al (2010). Some researchers promoting grammar practice (e.g., Bygate, 2001; DeKeyser (2007); Leaver, Rifkin & Shekhtman, 2004), questioned the traditional understanding of practice as mere repetitive action, and disassociated their understanding of practice from mechanical drills. DeKeyser, for example, puts particular emphases on the meaningfulness of practice and proposes the introduction of “loosely structured communicative activities” (DeKeyser, 2007: 6) for practising grammar, which follow explicit instruction and what he calls “anchoring exercises” (DeKeyser, 1998). Furthermore, Nassaji and Fotos (2011) dedicated their entire volume to exploring “communicative focus on form” (Nassaji and Fotos, 2011: vii) and various means of integrating a focus on grammar and a focus on communication in language teaching.

However, the majority of the above suggestions of integration appear to employ a needs-specific (thus unsystematic) principle, aiming to ““mop up” errors or gaps in learners’ knowledge” (Ur, 2011: 518). They present useful examples, inventive tools and effective techniques, but still no models or systematic frameworks based on integration, were suggested. In his critical review (2015), R. Ellis directly acknowledges that focus-on-form is better used “to refer to specific kinds of ‘activities’ or ‘procedures’ rather than to an ‘approach’”. In fact, the definition of grammar teaching itself has changed from “presenting and practicing” of grammatical features (Hedge, 2000) to a considerably broader and all-inclusive interpretation:

> *“any instructional technique that draws learners' attention to some specific grammatical form in such a way that it helps them either to understand it metalinguistically and/or process it in comprehension and/or production so that they can internalize it” (R. Ellis, 2006: 84)*

The above definition is clearly aiming to merge the explicit and implicit elements of grammar teaching but there is still an obvious dilemma of “and” or “or” to be addressed, as well as the ways of executing the suggested merge being not obvious. R.Ellis proposed a modular syllabus consisting of separate task-based and linguistic components (R. Ellis, 2003), which, although making a tentative step towards integration at a higher level, still kept the focus on form and the focus on meaning in separate sections of language curriculum.

Larsen-Freeman (2003) introduced a distinction between “meaning” and “usage” in language pedagogy, by proposing a theoretical framework of teaching grammar, which is based on three dimensions – form, meaning and usage, and is therefore called “three-dimensional”. She emphasised the dynamic nature of language learning, stating that “A difference in form

always spells a difference in meaning or use." (Larsen-Freeman, 2003: 44). To reflect that, Larsen-Freeman even introduced a term "grammaring", defined as "the ability to use grammar structures accurately, meaningfully, and appropriately", seeing grammar not as an area of knowledge but as "a fifth skill" "interconnected with the other skills" (Larsen-Freeman, 2003: 143). Nevertheless, she recognised that the three elements of her framework are different in nature and, for that reason, need to be learned differently. Unlike the integration proposed by R.Ellis (2003) above, where form and meaning were allocated to different sections of the curriculum, Larsen-Freeman suggested choosing activities for *each lesson* appropriately, depending on which of the three aspects the focus of learning was on, providing a wealth of ESL examples. The choice of how to implement these principles systematically within a curriculum though, is left entirely to teachers.

Meanwhile, SLA research has produced an abundance of empirical evidence confirming beneficial effects of Form Focused Instruction (FFI) on learners' language development (see reviews in Larsen-Freeman, 2015; Norris & Ortega, 2000; Spada & Lightbown, 2008), thus reinstating the necessity of teaching grammar in a language classroom, though the question of how to teach grammar effectively has remained open. Spada & Lightbown (2008) look at two ways of integrating FFI into CLT, introducing the terms of "isolated FFI", that is, FFI separate from communicative activities but incorporated into the CLT curriculum, and "integrated FFI", which is in line with Long's concept of short breaks dedicated to grammar *within* communicative activities, but includes proactive, as well as reactive, FFI (Spada & Lightbown, 2008: 186). They discuss different roles of isolated and integrated FFI, stating, however, that these are not competing against each other, but rather complementing each other. Moreover, the list of factors that can influence the choice of the type of FFI for a particular session, is presented. Their overall conclusion is that both types of instruction need to be included into a language curriculum, which is to be communicative, but the choice, similarly to Larsen-Freeman's, is needs-specific rather than systematic. That is why, I believe that the next step needs to be a design of a teaching framework where both approaches would be systematically integrated.

From a slightly different perspective, the discussion of possible ways of implementing different modes of introducing grammar features to learners has led to the emergence of the notions of deductive and inductive teaching. The former is understood as presenting grammar rules to learners explicitly before they use it in speech, while the latter implies creating the situations when the learners are exposed to the structures first, giving them a chance to work out the rules themselves (Haight et al., 2007: 289; also see a detailed review by Adair-Hauck, Donato and Cumo-Johanssen, 2005). Hulstijn (2005) considers both of

these as two versions of explicit instruction, as the rules are verbalised at some point during the lesson. The research comparing the two approaches has provided varied results (see review in Larsen-Freeman, 2015) and very rarely involved any kind of oral production testing. Similarly to other initially mutually exclusive teaching approaches, these were eventually combined by Herron and Tomasello (1992) in their "guided inductive model" using a technique that they proposed earlier and called a "garden path". They employed "hypothesis-testing procedures" but led their learners to the correct outcome by providing immediate feedback (hence the analogy with the "garden path") (Herron & Tomasello, 1992: 716).

A principle similar to the "garden path", appears to be at the basis of the PACE model, that was put forward by Adair-Hauck et al (2005) and was practically developed for teaching different levels in a language classroom. The researchers called to move beyond the opposition of explicit vs implicit teaching and aimed to "re-conceptualize grammar instruction" (Adair-Hauck et al., 2005: 274) by utilizing Vygotsky's principle of teaching within the Zone of Proximal Development (Vygotsky, 1996), which is discussed in later sections. They suggested:

- for the target structures to be presented (for "P") within a meaningful text (a story), rather than as isolated examples;
- then, for the attention (for "A") to be explicitly attracted to those structures;
- after that, for teachers and learners to "co-construct" (for "C") the understanding of the structures by engaging in a dialogue about them;
- and, finally, for learners to be provided with the opportunity to use the new structures in so-called "extension" activities (for "E"), which would take the form of a discussion evolving around the text presented at the start.

The model was trialled within a three month long project at a secondary school with rather positive feedback from learners. It could be possible to suggest that the PACE model attempted to practically apply Larsen-Freeman's theoretical three-dimensional model of teaching grammar, as Adair-Hauck et al integrated teaching of form, meaning and usage within one session. Haight, Herron and Cole (2007) empirically tested the effectiveness of the integrated model, based on PACE, applying it to teaching French in a second-semester classroom at a liberal arts college in the US. The PACE activities were slightly altered, as Haight et al used videos, rather than text as a basis for their discussions. The weekly video lessons were incorporated into the students' course schedule over the course of one semester. In their post-test the PACE students significantly outscored the traditionally taught groups, thus demonstrating that the PACE-type integration can be very effective for teaching

L2 grammar features. I can see this kind of integration of focus-on-form and focus-on meaning within one session, as a step in the direction of systematicity, which can be developed further to the level of a language curriculum.

Over the last two decades, the whole new field of academic inquiry - Instructional SLA (ISLA) - has developed within SLA, and is dedicated to investigating how language instruction could be adjusted, or "manipulated" (Loewen & Sato, 2017: 1-6) to make the learning and the acquisition of L2 more effective. This directly addresses the main aim of the present study. As ISLA researchers concentrate primarily on examining learning mechanisms and how the instruction affects them, the ISLA studies will be mainly discussed in the next section. Here, however, it appears appropriate to touch upon the ISLA discussion of the role in input and output in language teaching. On one hand, output, in terms of a classroom, is learners' production, which is a prime focus of the present study, investigating the effectiveness of inflection production. On the other hand, input, which is language material presented to learners, is the main source of grammar forms in the class.

In response to Krashen's Input Hypothesis (Krashen, 1982), stating that the language is acquired solely from the language input processed by a learner, Swain (1985) showed that comprehensible output has an important part to play in L2 acquisition. Similar to other dichotomous oppositions that were discussed earlier, each of these two points of view have gained a number of supporters among researchers, resulting in comprehension-based (CBI) and production-based instruction (PBI) directions in ISLA. A detailed analysis of these two approaches can be found in Shintani, Li and R. Ellis (2013), who analysed 35 comparative studies, conducted between 1991 and 2010. Their meta-analysis provided theoretical and empirical support for both types of instruction, concluding that both are beneficial for L2 learning, with no evidence for one being more superior, than the other. Thus, they could not recommend one over the other for more effective grammar teaching in an L2 classroom. However, they emphasised that there are different versions, methods and techniques within each of the approaches. For instance, PBI includes written, as well as oral production, with Grammar Translation and PPP given as examples. More importantly, though PBI is found to have been dominant in language teaching, it does not necessarily (and normally does not) involve learners' interaction, which is very important for the present study and which I will come back to in 3.3.4.

Following the above discussion, it could be possible to conclude that the integration of initially opposite approaches to teaching, such as "form-focused" and "meaning-focused" instruction, combining explicit and implicit elements of teaching have been producing quite

successful results and appear to be the most effective way of teaching grammar to L2 classroom learners. However, there seem to be an array of suggestions of how these elements could be integrated. Therefore, there appears to be a need for a systematic framework that would fully integrate the teaching of grammar form with the teaching of speaking and communication skills within a language curriculum, providing the opportunities for the wealth of the methods and techniques, which have been accumulated in language pedagogy research and practice, to be used systematically in teaching in the light of what we know about learning and acquisition. The creation of such a framework, which is also empirically supported, as well as being plausible within the constraints of language education, is one of the main aims of this study. Furthermore, Russian case morphology provides an excellent medium to examine the framework systematically, due to its complexity, which presents numerous acquisition, production and processing challenges to learners (see Chapter 4).

2.2. The learning of L2 grammar: explicit-implicit relationship

The efforts of language pedagogy researchers, as well as language teaching practitioners, to integrate explicit and implicit teaching of grammar, can be seen as parallel to the SLA debate on explicit and implicit learning of grammar. Initially, explicit learning was restricted to the classroom, while implicit learning was associated with naturalistic conditions and more or less equated to language acquisition (e.g., Krashen, 1981; Paradis, 1984). However, the concepts have evolved since and, at present, though there are still several definitions of these two constructs, the majority of researchers consider conscious awareness of what is being learnt, and often the intention to learn it (e.g., DeKeyser, 2003; Hulstijn, 2005; Kirkhart, 2001) as two essential components of explicit learning. The definitions of implicit learning focus solely on the processes that take place in learner's mind, for example, N. Ellis (2005) puts both instruction and the context of learning completely out of equation, focusing exclusively on communication:

> *"Implicit learning of language occurs during fluent comprehension and production. Explicit learning of language occurs in our conscious efforts to negotiate meaning and construct communication."* (N.Ellis, 2005: 306).

However, he suggested allocating explicit mode to initial stages of learning, which would, in his opinion, create the linguistic foundations for later implicit learning. This appears to be another attempt to bring both explicit and implicit elements into one language curriculum, while, at the same time, keeping them within different proficiency levels, thus separate. The idea of integration of explicit and implicit learning has been suggested (e.g., Rothman, & Slabakova, 2018), but the practicalities of it within the restrictions of a standard language

course are still to be addressed. That is why, by proposing my new integrative teaching framework, I plan to develop this idea further, particularly at the initial levels, when a lot of learning is traditionally explicit.

Hulstijn (2005) sees both explicit and implicit learning as “input processing”, differing in presence or absence of “conscious intention to find out whether the input information contains regularities” (Hulstijn, 2005: 131). In his opinion, explicit and implicit learning are different cognitive learning modes which can be activated for different tasks, possibly within one lesson. This allows me to infer that both modes could potentially be employed systematically within one grammar teaching framework.

The notion of intention of learning was reflected in differentiating between intentional and incidental learning. Hulstijn (2003) defined intentional learning as a mode when learners are informed, before engaging in a task, that they would be tested on particular aspect of its content. On the contrary, learning without this kind of advanced warning is referred to as incidental, with both definitions restricted to the classroom context. In real life though, the intention is not necessarily determined by the imminence of a test. According to Hulstijn (1989, 2005), the distinction between implicit learning and incidental learning is that the former is linked to the absence of explicit instruction on the language feature that has been learnt, while the latter – to the absence of intention on behalf of the learner. See more on this distinction in Denhovska, Serratrice and Payne (2016). The value of incidental learning has been also debated by Robinson (2005) and Schmidt (1990), who argued that incidental learning is highly unlikely for adult learners in relation to grammar forms which are not required for successful completion of the task. More recently, Hulstijn (2013) clarified that the term “incidental learning”, in the present research literature, is mainly referred to vocabulary learning, rather than grammar. Shintani and Ellis (2010) maintain the importance of incidental learning, interpreted as mastering grammar aspects while focusing on semantics, which seems to bring it closer to focus-on-meaning concept. However, they do not exclude the possibility of learners having some awareness of the grammatical feature, seeing this as distinct from implicit learning. As incidental learning cannot be systematised or built into classroom activities, this type of learning will not be discussed here any further.

Meanwhile, in psycholinguistics, the concepts of explicit and implicit learning have been connected to those of explicit and implicit knowledge. There is a wide range of definitions, as well as an academic debate around them (e.g., Bialystock, 1982; N.Ellis, 2005; R.Ellis, 1992, 1993, 2008). The former is often associated with being conscious and/or analysed, with learners being aware of them, while the latter is often said to be unconscious, unanalysed and/or with awareness absent (R.Ellis, 2008). Following this explication, explicit knowledge

is often equated to metalinguistic knowledge. For example, Hulstijn (2005) includes verbalizing “the regularities underlying the information one has knowledge of” into his definition of explicit (Hulstijn, 2005; 130). Pawlak, & Biedroń (2021) suggest that learners use explicit knowledge, when they have time, which is unlikely in communication, when the implicit knowledge is required (this distinction is discussed further in 3.2). In relation to the present study, the former can be concisely described as the knowledge of grammar rules, and the latter can be defined as the utilization of grammatical features in speech, or online real time production in speaking or writing. From the classroom point of view, despite the common agreement that the aim of both language teaching and language learning is primarily implicit knowledge, there still seems to be no consensus on how this goal could be achieved (see VanPatten & Williams, 2020, on the debate).

The discussion that stemmed from the above distinction evolved around possibility of transfer of explicit knowledge into implicit, in other words, whether the grammar that is learned consciously can lead to the unanalysed production of correct grammar features. This has grown into what appears to be one of the most prominent debates in SLA, often referred to as the interface debate, where “interface” is understood as a “connection or overlap between the two types of knowledge” (Han & Finneran, 2014: 371). There are three positions that have developed - non-interface, strong-interface and weak-interface (see detailed analysis and discussion in Han & Finneran, 2014), differing in the role assigned to consciousness in learning. This is very relevant to the present study investigating ways of effective integration of the two types of knowledge in a language classroom, namely, explicit knowledge of grammar rules and implicit production of grammar forms.

The non-interface researchers believe that explicit and implicit knowledge are formed and exist separately, which implies that rules that are learned in class do not convert into accurate L2 speech, thus consciousness plays no part in language acquisition. The researchers adhering to this position follow Krashen’s stance that learned knowledge and acquired knowledge are different (Krashen, 1981). They often refer to the finding by Paradis (2004) that implicit and explicit knowledge are associated with different areas in the brain, thus are different from the point of view of neuroscience. The crucial evidence pointing towards the absence of interface is fossilization (e.g., Hopp, 2010; Lardiere, 1998, 2005; Selinker, 1972), when some grammatical features are not acquired or acquired incompletely, despite the prolonged exposure to the L2 and increasing overall language proficiency, which implies that consciousness is unable to address this issue and, consequently, cannot influence grammar learning.

The weak-interface position is most clearly represented by the views of R. Ellis (2005) and N. Ellis (2005), both acknowledging some role of consciousness. R. Ellis believes that explicit knowledge can turn into implicit knowledge under certain conditions, for example, if learners are developmentally ready to acquire a certain grammatical feature. For N. Ellis, implicit learning and implicit knowledge are primary, but he acknowledges that explicit learning is able to assist in learning features that cannot be learnt or are difficult to learn implicitly. Thus weak interface supporters agree that explicit knowledge can be beneficial for language acquisition but assign different degree of importance to it.

The strong interface position sees language acquisition as a cognitive skill and, therefore, consciousness plays the crucial role in language learning, allowing any explicit knowledge to be converted into implicit (DeKeyser, 2007). This position is very close to the Skill Acquisition Theory, which is the realm of psychology and will be discussed in the next section.

Despite numerous difficulties in measuring implicit knowledge, Han and Finneran (2014) provide evidence that within one learner's language system all three interface options are possible. In similar terms, recently, DeKeyser (2015) promotes the idea of "synergy" of both types of learning, when one type would be more suitable for some grammar rules (mainly, simple and concrete), and the second type for others (normally, probabilistic). This is very important for the present study, as this allows me to infer that the potential grammar teaching framework would be most beneficial for learners if it systematically provides conditions where they could possibly exercise both explicit and implicit learning, thus optimizing the process of acquiring a language.

2.3. The acquisition of L2 grammar, and inflection in particular

One more concept which is central to this study is the acquisition of L2 grammar, as it relates to *processes* involved in developing L2 grammar by learners (Butler & Hakuta, 2004: 121). There appears to be a fuzzy distinction between acquisition and learning, for example, Krashen (1982) saw these as two completely different phenomena, restricting acquisition to naturalistic language environment and allocating learning solely to a classroom, while R.Ellis (1985) did not differentiate between these two concepts at all. Though Ortega (2014) states that in modern linguistics there is no such distinction and defines second language acquisition as a process of learning additional languages, the fields of language pedagogy, concerned with learning, and SLA, focusing on acquisition, have not merged yet, though the emergence of the Instructed SLA appears to be a step in that direction. The theoretical

distinctions between these two terms will not be discussed further here, as the aim of the present study is to analyse research findings from either of the fields, which could assist me in answering my Research Questions (see Chapter 5).

Similarly to the definitions of language learning discussed above, there are various explanations of how humans acquire languages. In this section, as well as giving a brief overview of main approaches to L2 acquisition, I will analyse how they view inflection acquisition in particular, which is the main focus of this thesis. In the present study, inflection, which belongs to a broader concept of functional morphology, is understood as "the expression of grammatical information through changes in word forms" that determine word functions (Baerman, 2015: 1-2) (in language pedagogy, inflection more or less equates to word suffixes). Despite being assigned a modest role in current SLA models, inflection has been identified as a major stumbling block in second language (L2) acquisition (e.g., Lardiere, 1998, 2005; Parodi et al., 2004; DeKeyser, 2005). Furthermore, it has also been recognized as "a prime candidate for fossilization" (Hopp, 2010: 902), a phenomenon when a particular grammar feature stops developing before reaching a target norm, despite the increasing proficiency level (Selinker & Lamendella, 1978: 187; also Selinker, 1972, Han, 2004; White, 2003).

The well-established generativist approach sees grammar forms, including inflection, as dependent on syntactic structures originating from the innate linguistic knowledge, described in terms of principles of Universal Grammar (UG). Initially, the L2 acquisition problems were interpreted by the restricted or absent accessibility of innate language learning mechanisms to L2 learners (e.g., Bley-Vroman et al., 1988; White, 1989; Meisel, 2000), as well as L1 transfer (e.g., Bellucci and Dal Pozzo, 2016; Brown and Iwasaki, 2013); both resulting in breaks in UG parameter re-setting (e.g., White, 1996; Schwartz and Sprouse, 1996). That, in turn, was assumed to lead to faulty mental representations of grammatical categories, known as representational deficit (e.g., Hawkins, 2003).

However, the small number of parameters did not account well for the immense variability in inflection acquisition, which is well-documented and empirically attested (see review in Stoll, 2015). The subsequent generativist research provided evidence that inflection deficit in production does not necessarily depend on the syntactic deficit. For example, Lardiere (1998), during her longitudinal case study of an L1 Chinese learner of L2 English, documented in detail the persistent absence of past and present verb inflection in subject's L2 speech, despite her perfect execution of verb movement, the syntactic procedure suggesting high level of L2 mental representations. An explanation was proposed by Prevost

and White (2000), who attributed the inconsistencies in L2 inflection production to the absence of surface realisation of correct grammatical categories (Prevost & White, 2000). In other words, it was assumed that the mental representations of language features were correct and the difficulties in inflection production were explained by the mapping of the lexical items onto the existing innate syntactic structures. This theory, called Missing Surface Inflection Hypothesis, is supported by the production data of two L1 Arabic learners of French and a Spanish and a Portuguese learners of German in the above study, examining verb tense inflection and its distribution in finite and non-finite contexts. Slabakova (2006) draws a similar conclusion that the stalling of L2 acquisition is caused by the imperfection of form-meaning mapping, negotiated by functional morphology, which includes inflection. According to her Bottleneck Hypothesis (Slabakova, 2006, 2008, 2014), functional morphology is harder to acquire than syntax, semantics or pragmatics, thus acting as a bottleneck for L2 acquisition. It is now put forward as "being at the core of language acquisition" (Slabakova, 2014: 7), with several experimental studies supporting the main postulate of Slabakova's hypothesis (e.g., Jensen et al., 2020).

The position of inflection within SLA has been upgraded further by the proposed "re-thinking" of parameter resetting as feature *(re)-assembly*, which can be formulated as producing combinations of morphological elements, specific for each language (Lardiere, 2008; also see Montrul & Yoon, 2009; and Slabakova, 2009; among others). In the same paper, Lardiere introduces the idea of Morphological Competence, which she describes as "the knowledge of precisely which forms 'go with' which features" (Lardiere, 2008: 109), acknowledging, though, that there is no concrete model yet that could explain this concept. However, issues of performance or processing are excluded from this notion, and no suggestions of how this new competence could be achieved, are offered.

The idea of Morphological Competence was further developed by McCarthy (2008), who suggested that mental representations are an essential part of the concept, at the same time acknowledging that the performance plays its part too. The two accounts of inflectional variability are proposed - "representational", linked to incorrect or absent representations of grammatical features; and "computational", connected to processing and performance (McCarthy, 2008: 460). She also suggests that syntactic deficit might not be directly connected with morphological deficit. As well as extending variability to comprehension, McCarthy's study develops further the Missing Surface Inflection Hypothesis, discussed above, by introducing the idea of a default value, which is not necessarily a zero inflection, which is illustrated by the analysis of L2 Dutch adjectival inflection, where a default value is an overtly inflected form (abundantly produced), which is opposed to the zero inflection in

neuter adjectives (which is substituted by an overt default). Furthermore, McCarthy demonstrated that inflectional variability is not only persistent but also systematic (rather than random), using the example of feminine adjectival inflection being regularly substituted by masculine against the absence of instances of opposite substitution. Finally, one of important conclusions made by McCarthy, is that inflectional variability extends to all proficiency levels, and can be different for different tasks. However, as Hopp has concisely summed it up, "neither the scope nor the causes of L2 inflectional variability have been conclusively identified" (Hopp, 2010: 902).

From the above analysis, it is possible to conclude that the generative hypotheses, which have been proposed so far, can contribute greatly to the description of learners' problems with regard to inflection morphology acquisition, but do not seem to be able to explain these. A few studies appear to point toward exploring inflection processing as the next possible step (e.g., McCarthy, 2008; Slabakova, 2009: 322;), as the complexity of online processing appears to provide a wide range of trajectories for L2 production, as well as accounting for learners' individual differences.

A number of alternative views of L2 acquisition are represented by the functionalist approach, which does not acknowledge the existence of innate linguistic structures in a human brain and places paramount significance on the relationships between form and function (e.g., Bates & MacWhinney, 1989) (or meaning, as in Janda, 2013), and inflection is seen as "an essential grammatical glue holding the relationships in constructions together" (Janda, 2007: 456). In a way, this echoes the earlier discussion of the relationships between meaning and form, as well as that of the role of explicit and implicit. Functionalists hold that linguistic structures, or representations, are derived, or "emerge" (Bybee & Hopper, 2001) from those very constructions, their "recurrent patterns and frequent use", observed in a language (Bischoff & Jany, 2013: 3). According to functionalists, form-function mappings form various connections, which are strengthened by frequent activations, that is why functionalist models often make use of connectionist neural networks and cognitive processing. Within this approach, language acquisition does not involve any inherent language-specific mechanisms. Instead, the purpose of a learner is to establish the target form-function relationships and realise them in communication (MacWhinney, 1997). MacWhinney's Competition Model (1997, 2005), for example, proposes that the L2 acquisition could be achieved by *restructuring* the existing connections and by creating new branches from existing networks, with the process mobilizing learners' cognitive resources and being aided by appropriate cues. The acquisition of inflection is not separated from the acquisition of lexis (Bybee & Hopper, 2001) and its deficiencies and shortcomings are

explained by the language processing constraints with too many competing factors (Menn et al., 2013). Thus, functionalists also arrive to the importance of learners' processing, as a crucial factor in the success of language production.

The literature review above allows me to draw out some rather common themes, present in two main SLA approaches with regard to grammar acquisition and L2 inflection in particular. Firstly, both see L2 acquisition as some kind of *(re)structuring* (or re-assembling) of the existing system of L1 or building an L2 system parallel to L1. This brings me to the concept of Interlanguage, which was first introduced by Selinker (1972) and described as an underlying system of an L2 language which is unique to an L2 learner and is different from the target language system, as well as from learner's L1. This gives a plausible explanation to the immense variability in inflection production, now commonly recognised and actively studied in SLA (e.g., Hopp, 2013, White 2003), which will also be investigated in the present study (Chapters 6 and 7).

Secondly, both generativists and functionalists assign an important role in L2 acquisition to mental *representations*, though the origins and the forms of these representation might be viewed rather differently. Though there is no clear explanation of a path which learners' brains take to form (or access) these representations, the insights into processing of linguistic information (that will be discussed in the next chapter), could possibly indicate some ways that could assist the formation of these representations.

Thirdly, both approaches appear to attribute difficulties in L2 inflection acquisition to *mapping* and *processing* challenges, though while for functionalists this principle underlies acquisition of language overall, for generativists it can be seen as restricted to functional morphology, because syntactic structures are assumed to be part of the language module in our brain. Fairly new body of the SLA research focusing on learners' processing, is emerging. For, example, Simpler Syntax theory suggests that "rules of grammar are taken to be pieces of structure stored in memory, which can be assembled online into larger structures." (Culicover & Jackendoff, 2006, 415). Finally, the investigation of the above processing issues is a fairly new direction and neither of the approaches seem yet to have been able to put forward any practical propositions with regard to the ways, in which L2 learners' processing difficulties could be effectively addressed. That is why, the present study pays particular attention to what is known about and can be investigated with regard to processes that take place in learners' brains during inflection acquisition and ways in which these could be made more effective; this is discussed in more detail in Chapter 3.

I am inclined to think that adult learners, at least at the initial stages of learning, synthesize information that they receive from various types of input, including explicit instruction, in order to establish the representation of the relevant target language structure. They can, then, access this information when they start producing the language, which, in turn, helps them further build their interlanguage - that is, their system of L2.

If the three themes above are now analysed from the point of view of language pedagogy and learning occurring in a classroom, the teaching of L2 inflection, firstly, needs to be organized in such a way that it would help a learner to *build their own mental system* of the target grammar (see Wong & Van Patten, 2003) (whether in the form of domain-specific mental structures or connection networks). This would require a systematic approach to the teaching of inflection. Secondly, the main concepts of the grammar required (e.g., grammatical categories, syntactic structures or word functions) need to be clearly and concisely formulated to enable learners, at the start of their learning, to form *simple representations* of these concepts in their explicit knowledge, to help them understand how they are related to each other and to provide a basis for the implicit L2 interlanguage system to develop alongside. (I accept that in the absence of naturalistic exposure to language and with maturational constraints of adult acquisition, these representations might be different from those of native speakers.) Thirdly, to enable *mapping* between these representations and real language forms, either produced or comprehended by learners, learners' processing needs ought to be considered; thus adding one more dimension to the potential teaching framework, which would be presented in Chapter 4. For this reason, the next chapter is dedicated to the analysis of relevant research related to cognitive processing, e.g., models of memory and skill acquisition, and how it affects learning and acquisition with the focus on inflection production.

Chapter 3
PROCESSING AND MAPPING OF INFLECTION

The large part of the discussion in Chapter 2 has pointed in the direction of form-meaning mapping and processing of form as the main difficulty in the acquisition of L2 grammar, including inflection production. That is why, in this section, I will turn towards analysing relevant research related to these issues as relevant to the research focus of this study – acquisition of L2 Russian morphology by beginner learners. As Skehan (1998) stated, "for a pedagogical intervention to be useful […], it must": realise what it is that learner is acquiring; describe "the minimal ingredients and mechanisms involved in language acquisition;" and understand how learners process those ingredients (Skehan, 1998: 48). I will look specifically at the actual mechanisms of processing of information in general and grammar in particular, and how these function with regard to explicit and implicit aspects of learning, attempting to dissect the process of acquisition into smaller components, in order to understand how those could be effectively assembled during language production.

3.1. Working Memory: encoding and retrieval of grammar forms during speech production

The discussion in Section 2.1 identified the need for integration of some dichotomous approaches to teaching and learning, as well as identifying a considerable body of research pointing towards learners' processing as the next direction of investigation. That is why, it appears to be important to examine mental processes that are essential for inflection production. As two influential figures in language learning research have put it:

> *"Understanding cognitive processes involved in L2 production is fundamental for determining if teaching for oral proficiency is in sync with the psycholinguistic reality of language processing in speech production. (N. Ellis & Larsen-Freeman, 2006)*

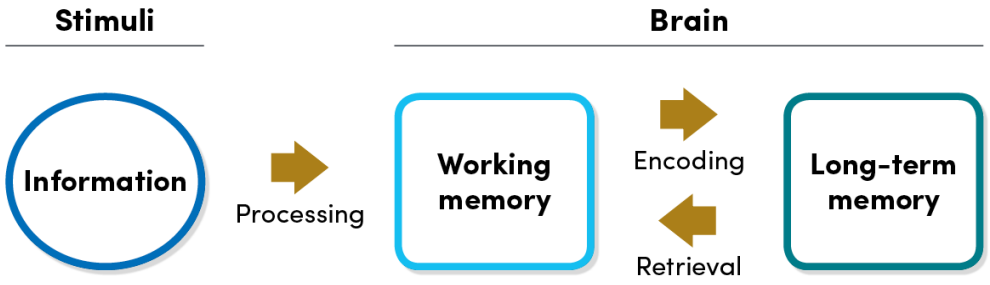
One of the significant breakthroughs in understanding the cognitive side of learning was the theory of Working Memory (WM), first proposed by Baddeley (1992), who defined WM as "the system for the temporary maintenance and manipulation of information" during processing (Baddeley, 1992: 281). Due to a large number of cognitive aspects involved in processing within WM, there have been suggestions to consider WM a part of cognition, rather than memory (see Juffs & Harrington, 2011). In the last few decades, WM attracted a lot of research in psychology and psycholinguistics (see ibid for an overview). Within the latter, there is a large body of research, which investigated the role of WM in language

learning and presented evidence of its direct involvement (see Mackey et al., 2002, for review), thus making it very relevant for the current study.

There is a wide range of definitions of WM, used in different WM theories (see Cowan, 2017, for a review), but the present study will be referring to WM, as described by Baddeley, as his theory is usually considered the most influential in the SLA research, as well as being identified as the most suitable for language production research (e.g., Kormos & Sáfár, 2006; Martin & N.Ellis, 2012). According to Baddeley, WM can be understood as a kind of buffer, where the incoming information (input) is stored temporarily, before it is passed on into the long-term memory (LTM) for further storing. This process is known as encoding. Moreover, when some information needs to be retrieved from the LTM, it is transferred back to the WM, before being produced (output) (see Figure 1); this process is referred to as retrieval (for more detailed summary of the elements of WM, see Juffs & Harrington, 2011). According to this view of WM, it serves as a buffer for transferring some input information into long-term memory and for retrieving relevant information from it, in order to develop a response. WM, thus, appears to be essential for storing and producing inflected forms, and consequently, is important for the present study.

Figure 1

Simplified Diagram of WM (Anderson, 2015)



In addition to the limited time, during which information is held in WM, the second important characteristic of WM is its very limited capacity. For example, Cowan (2017) proposed that it is capable of holding four “chunks” of information, though defining “chunks” seemed to be problematic. It has been shown that most humans can store about five words at one time or about seven digits (Anderson, 2015). Juffs and Harrington (2011) called WM “a bottle-neck through which information has to pass in order to be permanently stored in long-term memory” (Juffs & Harrington, 2011: 139). In addition, it has been demonstrated empirically that the size of WM varies between different individuals and is predetermined at birth but can

also be affected by other factors, for example age, physical or emotional state, as well as training activities (see review in Unsworth & Engle, 2007). Some empirical studies demonstrate that the size of WM can be a crucial predictor for the success of learning (e.g., Alloway, 2006), and that of grammatical knowledge more specifically (e.g., Fortkamp, 1999; Kormos, 2006), therefore, again, appears very relevant to the present study. WM is often referred to as one of the individual differences that moderate the effectiveness of L2 instruction (Pawlak, & Biedroń, 2021) and support the development of learner's interlanguage (Pili-Moss et al., 2020), thus are directly related to the overall aim of this study and the proposed teaching framework.

There is a reasonable body of research that investigates the correlations between WM and various aspects of language learning. Some empirical evidence links WM to successful performance in reading and to speaking fluency, though primarily in L1 (e.g., Daneman, 1981; Daneman & Carpenter, 1980; Nevo & Breznitz, 2014). Mackey et al (2010) provides a list of studies which showed the connection between WM capacity and the effectiveness of L2 input processing. However, it was found that the correlation between WM and different aspects of L2 speech production were not always consistent, especially in relation to fluency (Fortkamp, 1999, and Mizera, 2006). However, a comprehensive study by Martin and N. Ellis (2012) concludes that participants with higher WM capacity normally perform better in grammar-learning tasks. Moreover, they state that, in comparison with vocabulary, the processing of grammar patterns requires "more processing capacities, the holding of a greater amount of information over time, and the identification, selection, and correlation of relevant features both in the input and in long-term memory" (Martin & N.Ellis, 2012: 402), thus making WM of particular interest for my investigations. Finally, it appears that many studies investigating WM effects on grammar learning are conducted in laboratory conditions and there is a call for examining WM correlation with L2 grammar learning in a real classroom (Juffs & Harrington, 2011). The present study will be answering this call (see research design in Chapter 5).

Parallel to analysing the effects of WM and its correlations with learning outcomes, some cognitive psychology and psycholinguistics studies aim to reconstruct the processes inside a learner's brain, often involving WM, during the parsing of language forms. These studies often build on the SLA theories of acquisition and could potentially shed some light on the processes involved in learners' mastering of inflection. For example, Badecker & Kuminiak (2007) proposed their Working Memory Retrieval Model, examining language production. According to the latter model, the "elementary units", held in WM during production, are lexical representations with morphosyntactic tags (Badecker & Kuminiak, 2007: 68), thus

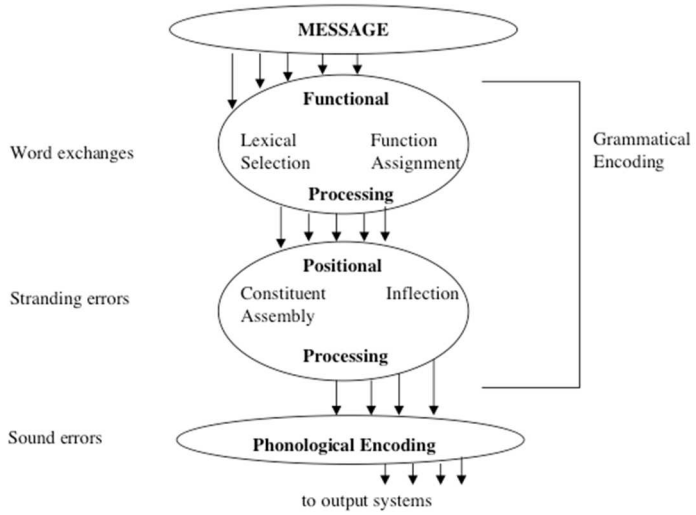
incorporating both meaning and form (see discussion in 2.1.). Therefore, the first cognitive step of production is described as “binding” selected lexical units with their grammatical and structural functions, with the aim of creating those “elementary units”. The speech production per se is seen as assembling them into various arrangements, assuming that the “elementary units” contain the information on how they could be combined at the sentence level.

This description of a Working Memory Retrieval Model echoes the discussion of mapping grammatical forms onto lexical units and (re-)assembling syntactic structures in 2.3. It is also in line with the widely accepted view of language production, which was developed (from Levelt’s (1989) earlier model) by Bock & Levelt (2002), who envisaged speech as consisting of two types of processes – “those that create a skeleton of an utterance and those that flesh the skeleton out” (Bock & Levelt, 2002: 405). They named the former as “grammatical encoding” and explained it as choosing appropriate lexical entries and using them to construct a sentence; the latter was called “phonological encoding” and referred to articulating and verbalising those sentences. Though Bock & Levelt (2002) used the term “encoding” in a different sense to Baddeley, Cowan and other WM researchers[1], their model describes widely acknowledged processes, which are assumed to be involved in the production of speech and thus, are important for me to analyse in relation to inflection production. Kormos (2014) believes that “understanding how one produces speech in an L2 is highly important in order to aid the teaching of this skill” (Kormos, 2014: xvii). Figure 2 shows that, according to Bock & Levelt’s (2002) model, after the message is conceptualized, the next stage is the selection of the right lexical item and identifying their function (Function assignment). Then follows the formulation stage, assigning of the right inflection and putting the constituents together (Constituent Assembly), thus, devising a mental plan for the utterance, before the actual articulatory output takes place (also, see Kormos, 2014, on speech production models).

Figure 2

An Overview of Language Production Processes (from Bock & Levelt, 1994: 406)

[1] McDonough (2005), in her description of Levelt’s model, addresses this discrepancy in terminology by referring to the two processes as “grammatical and phonological components” (McDonough, 2005: 82).



Badecker & Kuminiak (2007) developed Bock & Levelt's model further by analysing the cognitive processes in the context of WM, as well as using the standard cognitive psychology terms of encoding and retrieval, as defined by Baddeley (1992). They discussed the importance of WM for calculating grammatical agreements within sentences, as the items, which are to be agreed, need to be held in WM at the same time, thus increasing the demand on WM resources. The correct morphosyntactic information also needs to be retrieved from LTM into WM and assigned to the correct lexical item, therefore competing for the same WM space. Though the main aim of their investigation was gender agreement attraction in native Slovak speakers and the majority of the results of their three experiments do not appear relevant to the present study, their Working Memory Retrieval Model provides an up-to-date framework of grammar form production for my investigations. In addition, they examined various processing challenges that can lead to production errors, for example, competition between two sources of agreement, representational similarities (e.g., syncretic forms) and interference of different cues (Badecker & Kuminiak, 2007: 82), which are relevant for my analysis of inflection production difficulties.

It is worth noting that the above models were developed for the L1 speech. Kormos (2014) compares speech production in L1 and in L2, stating the common agreement that the three stages of conceptualization, formulation and articulation follow each other in this sequence (Kormos, 2014: xix) in both L1 and L2. She sees the main difference between L1 and L2 production in the processing that takes place during the formulation and articulation stages. (These correspond to Grammatical and Phonological Encoding in Bock & Levelt's model on Figure 2 (please note the difference in terminology discussed above).) While it is commonly assumed that formulation and articulation of an utterance in L1 are automatic, these stages

in L2 require cognitive resources and rely heavily on WM (also see de Bot, 1992; MacDonald, 2006).

Moreover, Kormos identifies two main production difficulties relating to these two stages: “a) resource deficits, b) processing time pressures” (Kormos, 2014: xxiv). She suggests that both of these difficulties can arise during both of the stages affected, namely, formulation and articulation. On one hand, resource deficits can manifest themselves during the retrieval of lexical units and morphosyntactic information, including inflection, (which would happen as part of formulating an utterance), as well as during the retrieval of phonological information, required for the actual verbalizing of the utterance (see Figure 2). In other words, learners might not know a word, a grammar rule or some pronunciation feature, which would be the consequence of the incomplete learner’s interlanguage (see discussion in 2.3) and is reduced with growing proficiency. On the other hand, the processing time pressures can cause faults in retrieval or no retrieval, as well as wrong articulation or no articulation, even if learners have the resources that they need, but their WM does not allow them to process those resources in time or to produce what they have processed. This view of processing difficulties in speech production offers very strong support for the Bottleneck Hypothesis (Slabakova, 2006, 2008, 2014) and Missing Surface Inflection Hypothesis (Prevost & White, 2000), discussed in 2.3. Also, the above discussion is paramount for my study, as the retrieval difficulties suggested by Kormos, apply directly to the production of inflection, which is part of both the formulation and the articulation stages.

Furthermore, according to Kormos, in L2 language production, the competition between the above retrieval processes for WM space, results in the processing of lexical, syntactic, morphological and phonological information being serial, that is each stage follows another, rather than running parallel, like in L1, this leading to slower and more erroneous production, even in advanced speakers (Kormos, 2014: xxiv-xxvi).

The order of retrieval of lexical items and morphosyntactic information is explained differently by different models, which are grouped into two types of theories, namely, spreading activation theories and modular theories (Kormos, 2014). In brief, the former assume that the structure of an utterance is formulated first and then is filled with lexical items, while the latter holds that the lexical items are selected first and then functional and structural tags are added later. The common knowledge that, in a classroom, learners traditionally start constructing an utterance with base forms of words, adding (if at all) inflection after the word is selected, appears to give support to the modular theory. However, helping learners to have at least a very basic syntactic structure in mind (say, Subject (base form) – Verb –

Case form), when constructing an utterance, from my experience, brings learners' attention to form from the onset. It might be possible that the retrieval of lexical items and grammatical forms during assembling an utterance is intermittent – the Subject is selected > its form is identified > the Verb is selected > the inflection is identified for it to agree with the Subject > the next lexical item is selected > the inflection is identified for it to agree with the Verb (also see Figure 2). In any case, whatever the order of lexical and grammatical retrieval during speech production, these processes clearly compete for limited WM resources and need to be optimized to produce a desirable outcome.

The main outcome of the above discussion is that it helped me to identify two main models explaining how grammar features are processed in our brain – the first one is general, the WM model by Baddeley (1992), showing how any information is processed in our memory; and the second one is specific to speech production, Bock & Levelt's (2002) model, demonstrating the processes involved in the oral production of grammar features, which would include inflection. It is apparent that the processes involved in formulating and producing grammatical forms are very complex, with the whole range of competing factors. In order to make it to the actual production, any grammar form, including inflection, has to go through a number of stages from initial encounter to the actual production, thus needs to be:

- registered in WM from the input;
- passed over to be encoded into the LTM;
- retained in the LTM, to be available for retrieval;
- retrieved for the purpose of production (during Function Assignment - Figure 2);
- attached to the relevant lexical item (during Constituent Assembly - Figure 2);
- articulated within the actual utterance.

All of these processes are essential for the successful production of any grammatical form, as the form normally has to pass through each of these to be produced. Though it does not appear plausible to investigate all of these in detail within one study, these stages will play an important part in my further discussion of the L2 learners processing of grammatical forms; their acquisition and possible ways of optimizing the process. At this point, it might be helpful to think of these as subprocesses of two main processes – the encoding (incorporating the first three, namely, registering grammar information in WM, passing it over into LTM and retaining it in the LTM) and retrieval (including the last three, namely, retrieving the encoded information back into WM, assembling the form and the actual oral production of the form). In 3.3., I will discuss how these two main processes could possibly be optimized in the language classroom, while, in the next section I will look at what research suggests

may happen with grammar information in our brain in between encoding and retrieval from the perspective of cognitive psychology.

3.2. Storage of grammar forms in LTM: two types of LTM and their role in language production

Having discussed the encoding and retrieval of information, as well as their subprocesses during speech production, I will now look closer at how research suggests that language information is stored and the relationship between the information that learners encode and what they then use during retrieval, especially in relation to the inflection production process.

Two principal concepts, which are considered to play a crucial role in acquiring skills, namely declarative and procedural knowledge, align directly with the concepts of explicit and implicit knowledge, discussed in Section 2.2. Their definitions were formalised by Milner et al (1998) as “knowing that” and “knowing how”, respectively (Milner et al., 1998: 449). Colman’s Dictionary of Psychology (2015) defines declarative knowledge (also often called explicit) as “awareness and understanding of factual information about the world” (Colman, 2015); and procedural knowledge (which is normally assumed to be implicit) as “information about how to carry out sequences of operations” (Colman, 2015). DeKeyser (2017) adds that declarative knowledge is “usually consciously accessible and often verbalizable, but not necessarily” (DeKeyser, 2017: 17).

These two kinds of knowledge are assumed to be stored by two separate memory systems, namely, declarative memory and procedural memory (both being part of LTM but located in different areas of the brain). Thus, any explicit explanations, for example, grammar rules, would be the realm of declarative memory, while skills, such as L1 pronunciation, would be under control of procedural memory. L2 inflection creates a considerable dilemma for psycholinguistics, as the rules about particular affixes appear to be declarative, while the fluent production of these affixes in speech needs to rely on procedural knowledge (Wright, 2018: 208). Considering evidence that low-level L2 learners’ knowledge is mainly of declarative nature (Ullman, 2005), it could be assumed that, for L2 instructed learners, this storage, at least initially, might be in declarative memory. It is particularly plausible in the context of classroom learning, as explanations of grammar rules are traditionally given in an explicit form. Kormos (2014) also suggests that speakers refer to their declarative knowledge while planning an utterance before the utterance is produced (see 3.1), while they are selecting a form (Kromos, 2014: xx). One more important difference between the two types of knowledge is that declarative knowledge can be easily updated, while procedural

knowledge takes some time and effort to acquire and even more so to change. This reinforces the need to invest into establishing strong inflection production skills from the start of learning.

There are subtle distinctions between the two pairs of the terms (declarative/procedural and explicit/implicit) from the point of view of neuroscience (see Dekeyser, 2017, for the explanation). Traditionally, declarative knowledge is assumed to be acquired explicitly, that is by being explained a rule or reading about a rule, while procedural knowledge originates from “doing”, thus is gained implicitly. However, this is open to debate (see DeKeyser, 2015, 2017). There is some research that shows that some declarative knowledge could be derived from implicit learning. However, Schmidt (1990) supports the opinion that declarative knowledge exists “on the continuum from unanalyzed to analyzed”, while procedural knowledge lies within a different “continuum from controlled to automatic” (Schmidt, 1990: 132). As an example, Kirkhart (2001) tested 60 undergraduate students on how accurately they could work out a rule by which 15 artificial letter strings were composed – half of the participants were told that they would need to work out the rules, while the other 30 were just asked to reproduce the novel strings. Though, the rules were never explicitly taught, Kirkhart found that both groups could formulate the rules, though the quality and the use of the declarative knowledge in the two groups was different. For instance, it was in direct correlation with the performance, as well as categorization, in the explicit group, but not in the implicit group. Also, the description of the rules in the explicit group were more abstract and generalized, rather than perceptual, which was the case in the implicit half. Consequently, Kirkhart differentiates between “abstract declarative knowledge indicative of the underlying rule structure for the explicit task and concrete declarative knowledge indicative of the physical dimensions of the stimuli for the implicit task.” (Kirkhart, 2001: 456). I believe that, for the production of correct inflection, it is the former type which would be more desirable, as it would be the abstract rules that would enable learners to formulate correct mental representations and build a solid interlanguage system. Therefore, explicit formulation of rules will need to be an essential part of the proposed framework (whether inductive or deductive). The distinctions between different dichotomies (declarative/ /procedural and explicit/implicit) are discussed by DeKeyser (2017), who, however, acknowledges that they are challenging to distinguish and are often not required for empirical research. That is why, for the purpose of this study, the above terms would be used as equivalent.

The relationships between the two kinds of knowledge (and between the corresponding types of memory) and their degree of influence on learning are not absolutely clear and are

the subject of on-going research. It has been suggested that learners' declarative knowledge has a potential of becoming procedural, thus implicit (e.g., see discussion of Han & Finneran, 2014, in 2.2; also Andringa & Rebuschat, 2015; Wright, 2018: 211). However, at present, the exact workings of the process, at least with relation to morphology learning and production, are not clear, especially how this transition could be negotiated by WM, if at all.

To start with, Skill Acquisition Theory (Anderson, 2015), widely applied by psychologists to various cognitive skills, offers a well attested explanation of how declarative knowledge can be "converted" into procedural. This process is called proceduralization of declarative knowledge and is reflected in three stages of skill acquisition: 1) *cognitive stage*, when learners encode declarative information about the skill via WM, and are assumed to form a declarative representation; 2) *associative stage*, when the declarative information is frequently retrieved into WM every time the action is performed, and, with the speed of retrieval gradually increasing, learners are forming a procedural representation of the skill, producing less errors and increasing the rate of performing the skill; and 3) *autonomous stage*, when the process becomes more automated, with less cognitive analysis, faster procedures and eventually no declarative representation required; thus, the process is automatized (Anderson, 2015: 211-212). DeKeyser, VanPatten, & Williams (2007) describe automatization as the process *following* proceduralization, rather than being part of it; this is later questioned by DeKeyser (2015). In any case, automatization is outside the scope of the present study, as my thesis focuses on the initial stages of proceduralization.

Anderson (2015) defines proceduralization as "the process of converting the deliberate use of declarative knowledge into pattern-driven application of procedural knowledge" (in other words, transforming explicit knowledge into implicit) (Anderson, 2015: 216). In his cigar-rolling experiment, Anderson demonstrated that the activation of the areas of the brain which are associated with declarative knowledge, decreases dramatically just after 5 days of practicing. It has been suggested that at the onset of proceduralization, learners use both declarative and procedural knowledge (DeKeyser, 2015), but with repetitive practice, say, switching gears or serving a tennis ball, the balance tips towards procedural knowledge, which is retrieved considerably faster. Further explanation is offered by the ACT-R model (Anderson & Lebiere, 1998), suggesting that, once the proceduralization occurs, the learner does not need to retrieve every section or item (e.g., 3 add 4 equals 7) involved in the process separately any more. It is assumed to be assembled into a chunk of procedural knowledge ("3 + 4 = 7") and is being retrieved as such, delivering the expected result in shorter times each time. This way, retrieval does not use all available WM resources (De Jong & Perfetti, 2011: 537) and has plausible potential to perhaps be able to eventually or

occasionally (depending on the skill) by-pass WM. It is possible that that is why the formulation stage in L1 is automatic, as it is likely not to involve WM (see Kormos, 2014, and discussion in 3.1). De Jong and Perfetti (2011) speculate that even complex language production rules could possibly be chunked and procedularized (De Jong & Perfetti, 2011: 538).

What is important to realize, as DeKeyser (2017) explains, is that proceduralization does *not* imply that the amount of declarative knowledge decreases or that declarative knowledge migrates to the area of the brain where procedural memory is located. Proceduralization is "forming procedural knowledge" (DeKeyser, 2017: 19) by creating and strengthening new neural paths, resulting in target-like behaviour. Thus, during proceduralization, learners are expected to go through stages "from mostly declarative to mostly procedural" (Dekayser, 2015: 105). In addition, de Bot (1996) argued that proceduralization is not linear and involves a lot of "tuning and restructuring", instigated by the "mismatch between the declarative and procedural knowledge" (de Bot, 1996: 547).

The application of Skill Acquisition Theory to language learning has sparked heated debates. On one hand, the majority of generativists (see 2.3.) disregarded "the wholesale application" of the theory mainly on the two grounds – the acquisition of language through communication (which does not require declarative knowledge) and a special place of language in SLA, which is seen as different from the rest of human cognition (see discussion in Wong and Van Patten, 2003: 416). On the other hand, there have been suggestions that proceduralization of language skills, as well as L2 grammar, could be achieved in stages, similar to those of other skills (e.g., De Jong & Perfetti, 2011; DeKeyser, 2015; Lyster & Sato, 2013; Rogers, 2011; Xie et al., 2020). There is no doubt, of course, that the patterns in producing, for example, nominal inflection involving phonological production, possible fusion of several categories (e.g., gender, case, number) and ever changing discourse contexts, are far more complex than those in, say, rolling cigars or doing arithmetic additions. However, this does not cancel a possibility of some principles of acquisition of some aspects of grammar being similar to those of other skills. Larsen-Freeman (2015) even called using grammar "the fifth skill", adding it to reading, writing, listening and speaking. In the present study, I will investigate how the effectiveness of language practice could be increased, in order to improve the inflection accuracy in oral production.

As a determined supporter of possible proceduralization of declarative grammar knowledge, DeKeyser advocates deliberate practice as a plausible way to proceduralized knowledge. Skehan (1998), who also supports practice as a way of developing fluency and automaticity,

suggests that frequent use of grammar features might be more “vital” for the acquisition of morphology than for other aspects of grammar, and considers practice for languages with rich morphology to be “all the more important” (Skehan, 1998: 18). DeKeyser (1997), in one of his first studies, conducted an experiment which demonstrated that the two explicit grammar rules could be proceduralized within 20 hours of learning. In his laboratory experiment, 61 University students, mainly undergraduates, were explicitly explained four grammar rules for an artificial language, which was specially created for this project, including 32 novel vocabulary items, and was called Autopract. Then, they had 22 hours (over 11 weeks) of computer assisted sessions to practise their comprehension and production of Autopract nominal and verbal inflection. During the experiment, participants’ were administered the same computer tests, matching the pictures to the aural sentences for comprehension and typing the suffixes in for the same pictures, as a form of production. As expected, the reaction time and the error rate decreased following the power function, which is found for proceduralization of other cognitive skills. This allowed DeKeyser to claim that declarative knowledge of grammar could be proceduralized in the same way as other skills.

However, in real classroom learning, unlike laboratory experiments focusing solely on a limited set of items, there are considerably more factors which are involved in proceduralization and which impact the process. Rogers (2011) measured the same characteristics of reaction time and error rate during verbal morphology comprehension and production by English-speaking learners, who have been studying Italian in real classroom environment at a North-American University, within a standard language course. He investigated the differences in participants’ performance on verb conjugation in present across three proficiency levels, having tested 30 beginners, 34 intermediate and 21 advanced learners. The testing consisted of picture identification (for comprehension) and picture description (for production) task, as well as preliminary testing of the knowledge of six verbs (from beginner level) that were used in comprehension and production tests. Rogers found that, in comprehension, the reaction time and the error rate increased significantly only from beginners to intermediate, while, in production, the speed and the accuracy improved following the power function between each of the three levels, which suggests proceduralization in line with the Skill Acquisition Theory. However, the study has raised a number of questions. Firstly, one of the acknowledged limitations of Roger’s study is the cross-sectional design of the skill acquisition testing (which is normally performed on the same group of learners), the results of which are, nevertheless, consistent with previous studies. Secondly, coefficient of variation within groups was calculated, which, with growing proficiency, decreased significantly for comprehension (as expected), but did not reach significance for production. Roger explained the latter by possible insufficiency of production

practice that students had had within their language courses, which is an important conclusion for the proposed teaching framework, suggesting that for effective proceduralization, abundant production needs to be part of language instruction. Thirdly, no discourse or full sentences were involved in testing, as only conjugated verb forms were required to be recognized or produced, thus limiting the generalizability of the results to real L2 oral production, which is investigated in the present study. Despite the limitations, the important value of Roger's study is in suggesting the possibility of proceduralization in a real classroom and that repetitive production is likely to contribute to it.

Importantly, in both DeKeyser's (1997) and Roger's (2011) studies, only familiar lexical items were tested – in the first experiment, the novel words that were inflected during the training sessions and, in the second study, the six most basic verbs which were conjugated by learners since the start of learning. However, the acquisition of inflection implies the application of rules to any items in the obligatory context, including those words which were not previously encountered or not inflected by a learner. This is a different level of use of knowledge, requiring abstract representations accurately formulated in learner's brain and accurately applied in new contexts (see 3.1 for discussion). This is the application that the current study is aiming to investigate.

If proceduralization is now viewed from the point of view of the two models that were discussed earlier in Section 3.1., namely, the WM model by Baddeley (1992) and the speech production model by Bock & Levelt (2002), the first stage (*cognitive*) would be argued to represent the encoding, that is, registering information in WM and then passing it on into declarative memory for storing. The second stage (*associative*) involves repetitive retrievals - that is, extracting the encoded (and stored) information from the declarative memory back into WM and using it for production; this somehow initiating the forming of the new procedural memory, thus crucial for acquiring a new skill. Following this logic brings me to a conclusion, that, for facilitating the proceduralization of a skill, both of these stages needs to be optimized, in order to ensure, firstly, firm encoding of explicit information into learners' declarative memory, and, secondly, multiple retrievals of that information back into WM during numerous attempts of performing the target action.

This is similar to the conclusion made at the end of 3.1., with two distinct differences, which are important for the present study. Firstly, the *cognitive* stage of the Skill Acquisition Theory can be assumed to rely solely on explicit information and does not include language input, considered essential for language acquisition (see discussion in 2.1. and further in 3.3.3.). Fully agreeing with the importance of explicit instruction for successful acquisition, I will

incorporate it into my proposed framework, but at the same time, I will analyse further the role of language input (see 3.3.3). Secondly, the *associative* stage introduces the idea of the repetition of information retrieval as being crucial for skill acquisition, which will be discussed in more detail in 3.3. However, there is one more difference, which is not as obvious, but, nevertheless, is not less important. From the six subprocesses, that I identified (using the two models) as being involved in the encoding and retrieval of an L2 grammatical form (see 3.1), Skill Acquisition Theory does not appear to attend to Function Assignment and possibly to Constituent Assembly (as in Bock & Levelt's (2002), see Figure 2). This level of complexity makes a crucial difference between the acquisition of grammatical form, including inflection, and that of other skills, as typically explained by Skill Acquisition Theory. For an L2 learner to produce a case form, they need to analyse the context of the utterance and assign a function to a lexical item they would want to use, using semantic and pragmatic categories. As I will show in 4.1, there is a large number of case functions, many of which are not clearly defined, which would considerably complicate or even disrupt the retrieval process (and possibly the encoding too). Furthermore, when (or if) the function is identified, there is often a number of suffixes which would be available for retrieval for the same case – this adds one more step in the process. Finally, once the correct suffix is retrieved, it needs to be attached to the target lexical item, which more often than not would be *different* from the one that the suffix was attached to during encoding. Thus what is to be produced is not completely what was encoded, as the case suffix would be only a part of the outcome (lexical item would need to be retrieved separately, possibly going through some similar stages – see 3.1. for discussion). This is why the present study investigates case production on new, as well as familiar items.

However, the above discussion does not cancel the idea of optimizing the *cognitive* stage (that is, encoding), as well as *associative* stage (that is, multiple retrieval), for effective acquisition, as discussed in 3.1. What needs to be considered is exactly what kind of linguistic information needs to be encoded and then retrieved, in order to account for Function Assignment and Constituent Assembly stages of production, as well as how that information needs to be organized for the most effective processing. Therefore, I will propose that grammar instruction needs to create opportunities for learners to gradually build a database of different contexts for particular grammar features (both linguistic and pragmatic), which they would encounter in the input and then access and replicate during their oral production, thus, learning to inflect a range of different lexical items within different contexts. (That is why I would prefer the term "multiple retrieval" to "repetitive retrieval", as it emphasizes numerous retrievals of, say, particular inflection, rather than repeating the same inflected form of a word.) This inference, I would argue, is also connected to the first

difference identified above, with regard to language input processing, which is not considered by Skill Acquisition Theory, but which would potentially provide the contexts required for encoding Function Assignment. This will be discussed in more detail in 3.3.3, 3.3.4 and in 4.1.2.

To sum up this section, I conclude that the stages suggested by Skill Acquisition Theory for the acquisition of cognitive skills needed for morphological production, align well with the main processes assumed by the WM model, namely, encoding and retrieval. I argued earlier that Skill Acquisition Theory does not explain certain processes of L2 grammar production, e.g. as proposed by the speech production model (by Bock and Levelt, 2002); this viewpoint, however, does not suggest that it is not relevant to language acquisition. It might be possible that the relationship between declarative and procedural knowledge could have several scenarios, similar to the three possible options of coexistence of explicit/implicit knowledge proposed by Han and Finneran's (discussed in 2.2) – full conversion of explicit knowledge into implicit, the conversion of some explicit knowledge and finally the complete absence of such. DeKeyser (2015) speculates that there could be various paths for proceduralisation for different aspects of language learning, with some aspects which might be likely or more likely to be proceduralized, some proceduralized faster than others, or those, which, unfortunately, fossilize or require cognitive effort every time, without ever being proceduralized (DeKeyser, 2015: 107). It is also likely that proceduralization might be different for different learners. Language acquisition is a complex multifaceted process and examining some possible ways in which learners' declarative knowledge could be potentially proceduralized more effectively is one of the objectives of this study and is discussed in subsequent sections. The proposed teaching framework aims to design language instruction with the skill acquisition stages, as well as the speech production stages in mind.

One other inference that can be drawn from the above discussion, which appears important to stipulate here, is that the proceduralization of grammar features could potentially be more effective at the start of learning. At the initial stages, main grammar concepts of the target language are introduced to L2 learners, with instruction focusing on one concept at a time, which would ensure more effective storing at cognitive stage. At the same time, some reasonable time is normally allocated in the curriculum for practicing specific grammar features, enabling multiple retrieval during the associative stage. That is why, in the present study, I investigate the acquisition of inflection in a classroom by complete beginners, who have no previous L2 knowledge.

3.3. Processing in a language classroom and possible ways of optimizing its effectiveness

In this section, I will discuss how the two main processes, involved in the production of L2 grammatical form, identified in 3.1 and 3.2, namely, the firm encoding of grammatical forms via WM and the retrieval of these forms during L2 speech production, could be optimized in the classroom learning environment. During this discussion, the two models, described in 3.1 and 3.2, will be referred to - the Working Memory model by Baddeley (1992) and the Speech Production Model by Bock and Levelt (2002) (explaining Function Assignment and Constituent Assembly stages, see Figure 2). Parallel to this, I will analyse which of these processes could be contributing to the proceduralization of the declarative knowledge that learners acquire about these forms.

3.3.1. Encoding and retrieval: which aspects could be optimized

Starting with encoding of explicit information, the limited capacity of WM (discussed in 3.1) is likely to be of great significance for the effectiveness of learners' processing in a classroom, where they have to encode and retain large amounts of explicitly explained rules and, then, utilize them in their production. Juffs and Harrington (2011) speculate about possible ways in which the characteristics of the WM could affect the success of learning. According to Juffs and Harrington, if WM is "a fixed trait", then this construct would not be useful to language pedagogy, but if the attentional resources within WM could be manipulated, then investigating the factors that facilitate this manipulation, would become of paramount importance for language learning (Juffs & Hurrington, 2011: 159). However, I believe that the limited capacity of learners' WM, even if it cannot be changed or changed significantly, could possibly be addressed by restructuring how the language material is presented and practiced in class, in order to make language instruction more processable. The main inference from this limitation is that if the information presented at a particular moment in class, cannot be registered in WM, it does not reach LTM in that instance; in other words, would get lost. Thus, it appears logical that explicit grammar information, presented to learners at one time, needs to be segmented and might be better introduced in small processable amounts, which their WM would be able to encode into LTM, (see Chapter 4 for further discussion). Nevertheless, there is still a question of the ways of how the explicit information encoded in LTM can be effectively stored and applied during language production, which is discussed further in this section.

In addition to explicit grammar rules, language learners process large amounts of other L2 information (while reading or listening in L2), defined as language *input*, that is "the language that learners are exposed to" (Gass, 2015: 183), which, among other things, contains grammar forms. VanPatten (2004) distinguishes between input and intake, which is defined as a "subset of the input that has been processed in the working memory", (VanPatten, 2004: 6). The challenge is how to increase the percentage of grammatical input, which would become intake to be stored in learner's LTM, which would be retrievable during production. This is discussed further in 3.3.2.

It has been demonstrated that large proportion of input is lost before the first retrieval (see Lightbown, 2008: 30) but numerous retrieval gradually increases the amount of information that is retained (for an example, see a summary of Ebbinghaus's forgetting curve in Schacter et al., 2009: 243 – 272). Several studies offer various mathematical schemes of these retrievals aiming to increase the amount of information firmly encoded in the LTM, as well as explaining the phenomenon from the point of view of neuroscience (e.g., Murre & Dros, 2015; Xie et al., 2020, Zhan et al., 2018). Though these studies reinforce the idea of practice, the majority of these experiments investigate memorization of information (often lists), using repeated learning procedure, that is, repeating the same material at different intervals. This kind of retrieval is often called passive in psychology (also "passive recall" or "passive review", or maintenance rehearsal), and has been proved not as effective for long-term retainment, as active retrieval (see Collins, 2019, for the overview). In language learning, this kind of retrieval of explicit grammar rules (in the form of rereading case paradigms or repeating the same phrases after an audio) has been deemed ineffective for grammar acquisition (see 2.1.; also discussions in Janda & Tyers, 2021; Robinson, 2008). It could be speculated that when paradigms are rehearsed, the Function Assignment (discussed in 3.1 and 3.2) is not activated (as there is no context), and neither is the Constituent Assembly (as typically ready-made forms are presented). Rehearsing chunks after an audio appears to exclude these two stages too. For these reasons, in the present study, I will focus on active types of retrieval. It is important to note that the above discussion does not cancel the importance of multiple retrieval for the effective storing of the encoded information (see 3.2) – it just specifies a more effective type of retrieval.

Furthermore, when learners deliberately retrieve the encoded information back into WM with the aim of using it in their production, the crucial question is the type of classroom practice which would encourage numerous retrieval instances (see 3.2.). This brings up the concept of practice in language learning, which was discussed in 2.1., and whose definition in language pedagogy or SLA is far from being straightforward (see Larsen-Freeman, 2005).

As discussed in 3.2, cognitive psychology sees practice as a repetition of attempts aiming to match learners' actions to the instructions (Anderson, 2015), which is a type of repetitive retrieval. For language learning, however, that does not appear to be sufficient, as retrieving the same form is not enough, as, to be able to speak, learners need to apply their declarative grammar knowledge to *new* contexts, rather than performing the same action in the same circumstances. In language pedagogy, it has been demonstrated over the years that rote repetition or grammatical drills did not deliver the desired implicit knowledge or expected levels of acquisition (see discussion in Wong & VanPatten, 2003). A more recent definition of practice by DeKeyser as "specific activities in the second language, engaged in systematically, deliberately, with the goal of developing knowledge of and skills in the second language" (DeKeyser, 2015: 1), though suggesting systematicity and consciousness, appears to be fairly general and could potentially relate to rather different kinds of activities. Here, I will attempt to narrow these down by analysing mental processes involved in different types of retrieval and some factors that make the retrieval of grammatical form more successful.

In cognitive psychology, two types of active retrieval are typically distinguished, namely, recognition and recall (e.g., Bower, 2000). The former is normally understood as the activation of the information stored in declarative memory triggered by a relevant cue in the input, for example, when a particular word or a case suffix that has been studied, is identified in a text or an audio. The latter can be defined as the deliberate accessing of information stored in LTM, without external cues, and can be illustrated by a voluntary use of a word or a case suffix in an appropriate context, when speaking (see detailed descriptions in Gillund & Shiffrin, 1984). A more recent simpler model offers to see "recall as simply retrieving items when probed with a cue […] and recognition as retrieving cues when probed with items" (Srivastava & Vul, 2017: 292). The latter model explains these two processes in terms which appear more appropriate for the learning of inflection, as explicit grammar information about, say case functions and case forms, could serve as cues which are activated during the recognition, and which then initiate the production of grammar forms during recall. What is also important is the conditions that would make both processes more effective in the context of grammar learning, making the required cues more distinct, in other words, kinds of classroom activities that would be able to optimize these.

Therefore, in Chapter 4, I will propose a teaching framework which would enable learners to process the required L2 information by managing the resources they have in the more optimal way, thus hopefully reducing variability. I argue that, in order to make the encoding of the explicit grammar information more effective, it needs to be presented in smaller

processable amounts, for it to be registered by learners' WM and then to be encoded in learners' LTM. The ways of optimizing the two types of active retrieval, namely recognition and recall, in the context of grammar learning, will be discussed in the next two sections and will become an essential part of the proposed framework.

3.3.2. Recognition (as a type of retrieval) and input processing

Now, I will analyse the role of recognition, the first type of active retrieval, in the effectiveness of language learning, which was first demonstrated by Schmidt (1990, 2001), who coined the term 'noticing' and conceptualized it as a deliberate attention to linguistic form within communicative activities (Schmidt, 1990). Schmidt discovered that, when he was learning Brazilian Portuguese, he only produced inflected forms which were present in the recorded input. What is important, not all forms which were used by Schmidt's interlocutors, were present in his speech, but only those that he 'noticed' and noted in his journal. Schmidt considered 'noticing' "the necessary and sufficient condition for converting input to intake" (Schmidt, 1990: 129), which is the question that has been debated with regard to the role of WM (see 3.1). This claim is generally known as Noticing Hypothesis and has been actively discussed in both language pedagogy and SLA, since its first introduction. It is considered a different process from comprehension, as the latter focuses on meaning and semantics, leaving "the underlying interlanguage system untouched and unscathed" (Skehan, 1998: 15).

Robinson (1995) examined the exact cognitive mechanisms involved in 'noticing', considering various attention and memory theories. He described 'noticing' as "a consequence of encoding in short-term memory", with the result being "subsequently encoded in long-term episodic memory" (Robinson, 1995: 298). He also defined it as "detection with awareness and rehearsal in short-term memory" (Robinson, 1995: 318; also see Izumi & Bigelow, 2000 for review). According to Robinson, the effectiveness of encoding in LTM depends on the level of activation that occurs in WM every time 'noticing' happens, as well as on repetitive rehearsal. This makes 'noticing' a type of multiple retrieval (recognition in this case) of the information encoded in the declarative memory, thus strengthening the neural retrieval paths, which is suggested as a possible way to proceduralization (see discussion in 3.2). Unlike in Ebinghaus's hypothesis, when the same information was retrieved passively again and again, when 'noticing' takes place, the same grammar information (or cue) is retrieved (recognized) actively. For example, rather than retrieving exactly the same inflected forms in a list (as in passive retrieval), the same inflection, during 'noticing', is retrieved in different contexts, as well as added to different

lexical items, thus establishing Function Assignment for particular grammatical forms. That is why, Schmidt's finding, which might be modest from the point of view of SLA, appears of considerable importance for language pedagogy.

Furthermore, VanPatten (2004), in his seminal work on the Theory of Input Processing, views 'noticing' as the first step leading to input processing, which he describes as "making form-meaning/function connection in real time comprehension" (VanPatten, 2004: 7; also VanPatten, 2015). He understands it as the process that follows initial 'noticing', when L2 learners establish the actual connections between form and meaning, in other words, retrieve grammatical cues. This loops back to form-meaning-function relationship, which was discussed in 2.1 and makes input processing directly related to the teaching of inflection. N. Ellis (2005) links 'noticing' directly to the reinforcing of form-meaning mappings.

Moreover, VanPatten (2004) considers 'noticing' *prerequisite* for the language acquisition of a grammatical form. He discusses Schmidt's (1983) case study of Wes, a Japanese learner of English in Hawaii, who could communicate without ever acquiring linguistic form because he never noticed any inflection in the input, despite having plenty of it. That meant that recognition of grammatical form did not take place, and raised the question about conditions, under which 'noticing' occurs. Schmidt suggested that learners' attention needs to be explicitly drawn to new grammar forms and saw explicit instruction as a means raising learners' awareness of grammatical form, thus forming the expectation required for the cognitive processing of new grammar features in the input and developing their interpretability. It is interesting that, in psychology, some recognition models suggest that recognition is only successful, if the participant chooses to perform a memory search (Gillund & Shiffrin, 1984). This reinforces the role of consciousness in 'noticing' a grammar form, as a deliberate search with the aim of retrieving linguistic cues, as in Srivastava & Vul (2017) (see 3.3.1).

Van Patten (2020) comes to the same conclusion about the need of conscious awareness but explains Wes's phenomenon by the primacy of lexical processing during comprehension of input, that is, L2 learners process lexical items *before* grammatical form. He concludes that 'noticing' is restricted by the WM capacity, as learners, focusing on meaning, have no cognitive resources left for noticing the form, unless specifically instructed. That is why, according to VanPatten (2004), 'noticing' functions as that specific instruction, encouraging learners to process grammar forms in the input, in order to then pass this information into their memory for storing.

To put his Theory of Input Processing into practice, VanPatten has developed Processing Instruction, which is defined as "an instructional intervention that seeks to alter certain faulty processing strategies that language learners exhibit" in processing L2 input (Comer & deBenedette, 2011: 647). The three corner stones of Processing Instruction are: 1) explicit information about the target structure; 2) explicit information about processing strategies; and 3) structured input activities (Wong, 2004: 36). This further emphasizes the role of conscious awareness of L2 form as prerequisite of its acquisition. My proposed teaching framework will adopt this assumption and explicit instruction will be an essential part of it (see 4.2.2). Though modifying the L2 input, in order to encourage learners to process grammatical features, as an essential element of Processing Instruction, has been demonstrated to increase learners' understanding of grammar forms and in some cases even their production (e.g., Comer & deBenedette, 2011), I believe that systematic modification of input might not be plausible within the time constraints of a standard beginners' course. That is why, that part of Processing Instruction will not be part of my proposed teaching framework, while explicit instruction and encouraging learners' 'noticing' appear to have more potential in this respect. The explicit explanations are already part of any L2 grammar syllabus, while 'noticing' would not be time consuming or require extensive preparation. In addition, VanPatten's suggestion of primacy of lexical processing by itself is also an important inference for the proposed grammar teaching framework, as it implies that, in order to ease learners' cognitive load, teaching new grammar should be carried out exclusively on familiar vocabulary.

From a slightly different perspective, retrieving relevant linguistic cues, during recognition, links a grammar form to the initial grammar concept of, say, a particular case, which was explained explicitly. This way, 'noticing' can be assumed to encourage deeper levels of processing of the L2 input, which involves thinking about conceptual meanings of items and is well-attested in research as a facilitative condition for successful storage in LTM and for the accuracy of retrieval (see overview in Gallo et al., 2008; also Lightbown, 2008; and McLeod, 2007). Robinson (1995) went further, differentiating between "data-driven processing" and "conceptually-driven processing" which activates "a schemata or higher-order relations" from LTM, which, in turn, lead to organizing items noticed in the input, into more abstract configurations (Robinson, 1995: 299). The latter could possibly be linked to Kirkhart's (2001) (in 3.2.) "abstract declarative knowledge" and be interpreted as forming mental representations of grammar features, which are thought to enable learners to apply grammar rules to new contexts and are considered essential for L2 grammar acquisition (see 2.3). In line with this, VanPatten suggests that 'noticing' could possibly be a first step towards "accommodating" of a grammatical form within learner's interlanguage and trigger

its restructuring (VanPatten, 2004: 25). Thus, it could be inferred that 'noticing' makes input processing the first essential step in language acquisition.

The facilitative effect of 'noticing' within classroom instruction has been advanced by other researchers (see review in Izumi & Bigelow, 2000; also, R. Ellis, 1993; N. Ellis, 2005; Fotos, 2001; Izumi et al., 1999) and empirically supported since (e.g., Brown, 2016; Gurzynski-Weiss & Baralt, 2015; Mackey et al., 2002). In Mochizuki & Ortega's (2008) experiment discussed in 2.1, participants' performance on relative clauses structure significantly improved after their experimental group received, what they called, "grammatical guidance to attend to a specific L2 form" (Mochizuki & Ortega, 2008: 11), which would, in fact, encourage 'noticing', and consequently enhance the retrieval of previously encoded information. In a way, 'noticing' echoes Focus-on-form, which was discussed in 2.1 and was actively promoted by Ellis (2003) and Larsen Freeman (2015) in their "focused tasks" (e.g., Larsen Freeman, 2015: 269), without calling the term 'noticing' though. However, there do not appear to be any specific guidelines offered how noticing could be systematically facilitated within classroom activities.

In view of the above, if 'noticing' is actively encouraged and facilitated within classroom activities, it could potentially be one of the ways of manipulating the attentional resources that are often associated with WM (see 3.1) during input processing. Juffs and Harrington (2011) suggest that learners "could be assisted in focusing on L2 form", and "learning to control their attentional resources […] could be a key to more successful L2 learning" (Juffs & Harrington, 2011: 159).

As discussed above, in a classroom, learners process two different kinds of information related to L2 grammar – explicit grammar explanations and grammar forms in the L2 input (when reading or listening). In order to enable the effective processing of the former, the explicit information needs to be restructured to address learners' processing requirements, as discussed in 3.3.1. To optimize the processing of the latter, I propose that 'noticing' needs to be encouraged during reading and listening, to facilitate repetitive recognition of grammar features in the meaningful input. It appears to be one of the means of ensuring that the inflection form and meaning are connected and are securely kept in LTM. For this reason, the proposed grammar teaching framework will incorporate stimulating 'noticing' as an essential part of the instruction.

Furthermore, the research above points that once grammar forms are 'noticed' in comprehension of the input, they eventually start appearing in production. This allows me to

infer that this firmly encoded information also becomes more easily available for further retrievals, thus establishing a mutually beneficial relationship between encoding and retrieval. Then, it is not unreasonable to think that attracting learners' attention to grammatical form could potentially assist in Function Assignment during the later production of grammar forms (see 3.1). Moreover, it could be logical to suggest that emphasising and encouraging 'noticing' within classroom activities, could potentially, strengthen the connection between the declarative knowledge about, say, inflection, and its production in speech, thus possibly making its contribution towards proceduralization. Further speculation about whether this kind of cognitive activity leads to creating stronger mental representations of the feature in question, though open for debate, might have some ground. In any case, the well-attested role of 'noticing' in facilitating language learning through multiple recognition, secures its place within the proposed grammar teaching framework.

3.3.3. Recall (as a type of retrieval) and the Output Hypothesis

Now I will analyse the role of the second type of active retrieval, namely, recall, in the acquisition of L2 grammar features. Unlike recognition, during recall, grammar forms and grammar functions are retrieved from LTM, "probed by the cues" (Srivastava & Vul, 2017) contained in the message context (see 3.3.1). As discussed in 3.1., an L2 learner has to assign a function to the target lexical item, then recall the correct form (e.g., inflection) and, after that, articulate it. In L2 learning, the language which is produced by a learner is called output, and its role in L2 acquisition has been debated in SLA for a number of decades.

In her Output Hypothesis (see discussion in 2.1), Swain (1985) argued that L2 output is not just the end product but a means of language development. She recorded a day of lessons in one of the French immersions programmes in a Canadian school, and realised that very little learners' production occur in language classes. That prompted her to suggest that low performance of the students in speaking could be explained by no opportunities to produce language during their learning, and allowed her to conclude that in order to learn to speak, learners need to speak. Later, she suggested a few roles that L2 output plays in developing learners' interlanguage, – 1) to trigger 'noticing'; 2) to test hypotheses about linguistic forms and functions; and d) to develop metalinguistic knowledge (Swain, 1995: 128). (The fourth function, namely, to enhance fluency through practice, is not within the scope of the present study.) These functions of output were reiterated and developed by Gass (2015), Kees de Bot (1996) and Wong & Van Patten (2003), among others.

The first role/function acknowledges that, during L2 production, learners do notice the gaps in their L2 system or discrepancies between the forms that they produce and the target L2 forms. The second role implies that 'noticing' "pushes" learners to explore the options of how to fill these gaps, thus instigating cognitive processes and engaging them in linguistic analysis. According to Swain, output forces "learners to move from the semantic, open-ended, nondeterministic, strategic processing prevalent in comprehension to the complete grammatical processing needed for accurate production" (Swain, 1995, 128). This, in turn, either consolidates previous knowledge or instigates the creation of new knowledge, which constitutes the third role (Swain & Lapkin, 1995: 384). Gass (2015) refers to "pushed" output as "a subset of output" involving conscious effort. This, in a way, can be thought as similar to how intake is considered by VanPatten (2003) (see 3.3.2) a subset of input that is encoded in memory, often requiring conscious effort of 'noticing'. Swain's (1995) conclusion that "output thus would seem to have a potentially significant role in the development of syntax and morphology" (Swain, 1995: 128) underpins my proposed teaching framework. Larsen-Freeman (2003) emphasized the role of "pushed" output by saying that "pushing learners to express themselves clearly would be beneficial in that it would mean that learners would have to learn to process language syntactically in addition to processing it for meaning" (Larsen-Freeman, 2003: 103).

Swain and Lapkin (1995) conducted a study when they asked 18 13-year-olds of different proficiency levels from the French immersion programme, to think aloud while writing an article in L2 and then again when they were editing it. They demonstrated that L2 learners exercise 'noticing' without any outside stimulus and the majority of instances lead to the successful repair of the noticed gaps. Moreover, more grammar-related instances of 'noticing' (as opposed to semantic-related) were registered during the editing stage; these resulted in the changes in learners' metalinguistic knowledge, as registered by the participants' think-aloud protocols. Swain (1995) posited that, during production, the attention to form is stimulated by the perspective of the need to produce that form; the view also supported by other researchers, e.g., Izumi and Bigelow (2000), de Bot (1996). In addition, Swain and Lapkin (1995) suggested that there could be two possible "triggers" for output-related 'noticing' – external feedback, that is the input (that would represent recognition) or comments from the interlocutor, and internal feedback, when a learner notices their linguistic errors or difficulties in production in the absence of the former, utilizing their own cognitive resources (that is, during the recall).

However, output can have various forms and, while Swain & Lapkin's study investigated writing, speaking is considered by far a more complex process (Swain, 1995). Oral

production increases the demand on cognitive resources by having the additional stage of articulating the planned and formulated utterance (see discussion in 3.1), as well as by the constraints of the online processing, thus being considerably more challenging for L2 learners. At the same time, it appears to instigate the processes that benefit language development.

Kees de Bot (1996) analysed the cognitive processes that are potentially involved in the oral production of the L2 output, using the model by Levelt (1989) (see 3.1), who considered language production a type of information processing. He saw the value of output in activating the information stored in memory, and strengthening of the connection between that information and the language forms that are produced, without analysing the exact stages of speech production (which will be discussed again later). Furthermore, he suggested that if the connection is wrong and there is some discrepancy between what is produced and what is encoded in memory, then the error message (feedback) is sent back to the brain, which would stop this connection from being used again, in other words, would prevent the faulty retrieval from being executed again. This is how he envisaged 'noticing' functioned during language production. I can see this view supporting the idea of multiple retrieval of the encoded information eventually leading to correct production (see 3.1 and above). According to de Bot, 'noticing', by itself, does not solve the problem of a gap in production, but attracts learner's attention to the required information, thus initiates the solution. Izumi and Bigelow (2000) suggest that 'noticing' during production, "stretches" learners' interlanguage "capability", forcing L2 speakers to modify their output (Izumi & Bigelow, 2000: 243). Mackey et al (2010) speculate that, during the modification of output, "more elaborated processing of linguistic information" might take place, when learners process grammar that they have encoded but have not used in production (that is, not recalled) yet. In addition, they suggest that noticing a gap in production encourages learners to identify the place for the grammatical feature in question in their L2 system (Mackey et al., 2010: 504); this could be linked to forming L2 representations.

Moreover, de Bot (1996) noted that output per se does not lead to the creation of new declarative knowledge (unless a textbook or a dictionary is consulted) and proposed that the importance of it, is in the "transition" from declarative knowledge to procedural (de Bot, 1996: 549). Developing this further, he defined acquisition "as gradual growth of knowledge structures and an increase of the ease with which those structures can be used in processing" (de Bot, 1996: 531). This view of acquisition appears quite relevant to the present study.

It is important to mention that Output Hypothesis never aimed to contradict or oppose the Input Hypothesis (see 3.1) but rather complement the latter. Izumi et al (1999) and Izumi & Bigelow (2000) conducted two experiments analysing the effects of combining input with two different types of output, namely, essay writing and text reconstruction, on the acquisition of English conditional. These two treatments were administered to the experimental group of adult English language learners in different order, to address the task effects. After each task, participants were provided with some written input, where they were asked to underline features which they thought were most important for their future production – that was used as a measure of 'noticing' (thus, 'noticing' was not measured in production). The control group were asked to underline features which they thought were important for comprehension. Though the researchers did not get the statistically significant difference between the experimental and control groups, the results demonstrated that learners who produced oral output, showed significant gains on the majority of seven features identified for the English conditional, while, in the control group, the significant improvement only occurred on one of the features. Izumi and colleagues considered this indicative of the advantage of the oral output group. The absence of the significant difference was explained by a fairly high level of accuracy in the pre-tests. It feels appropriate to note here that neither of the studies presented any explicit instruction, the importance of which was emphasized by VanPatten (2004), which might have also contributed to low levels of 'noticing'. Though the results of these two studies did not confirm the three hypotheses that the oral production group would perform significantly better than the control group, the researchers made a start in the empirical investigations of the role of the integration of both input and output in grammar acquisition and laid the foundations for further investigations. The most important result of these studies is that the significant improvement of the participants' performance was only registered after the second treatment (in both studies). This implies that, for increasing the effectiveness of acquisition, both input and output processing need to be facilitated in language instruction, embedding a short-term output treatment in every task. This is one of the aims of the proposed teaching framework. In addition, both studies recorded greater within-group variation in performance, as well as less 'noticing', in the essay-writing task in comparison with the text reconstruction treatment, which allowed them to conclude that 'noticing' is more effective in oral production. This last conclusion reinforced my decision to investigate speaking in the present study. Finally, the production investigated by Izumi and colleagues, did not involve interaction, which is discussed further.

3.3.4. Recall and the Interactive Approach

In this section, I will carry on discussing the ways of increasing the effectiveness of recall of grammar forms but now in the context of interaction, rather than one-way production.

The ways in which oral output, specifically in communication, could influence learners' language development were examined by Skehan (1998), who expanded the four roles of output, proposed by Swain (1995), to six functions of output, adding the development of discourse skills and personal voice. Though I will not discuss the added functions here (as they are related to communicative competence, which is not the subject of the present study), Skehan's analysis of the impact of communication on learners' interlanguage, appears of great interest. In addition, he included morphology into "Systemic knowledge" essential for communication, together with syntax and semantics (Skehan, 1998: 14), which is an important step forward in acknowledging the importance of form in the construction of meaning. Though Skehan did not discuss 'noticing' as such, he paid a lot of attention to "syntactic processing" (as a function of output), which is assumed to be initiated during output production and has "beneficial spin-off effects on underlying interlanguage" (Skehan, 1998: 20). He drew upon two previously proposed frameworks for repair of communication breakdown. One of them, the negotiation-of-meaning approach suggests that during communication, learners identify the limitations of their interlanguage, which cause difficulties in conveying their message, and overcome them by using the feedback provided and consequently encode new meanings; this, in turn, leading to the changes in their internal L2 system (Skehan, 1998: 20). At a closer look, the first part of the described process is very similar to the triggering of 'noticing', proposed by Swain (1985, 1995) (see above), while the end part could possibly be interpreted as forming new mental representations, required for grammar acquisition, which is an important implication for my proposed teaching framework. The other framework that Skehan discussed, was that for "strategic competence" which is seen as a more general "improvisatory manner" to resolve breakdowns (Skehan, 1998: 21), in a way similar to the third option, described by McDonough (2005) (see below).

However, Skehan points out that the changes to learners' interlanguage are subject to a number of external factors and formulates three conditions which are essential for those changes to occur. According to Skehan, for the improvisatory repair or communication breakdown solution to result in the interlanguage modification, it - firstly, has to leave "a trace" in learner's memory, that is, to be "noticed", which results from the interaction being "sufficiently salient and/or the processing capacity available"; secondly, must be generalizable, that is have the potential to be used again; and, thirdly, needs to be proceduralized, either through a strong impact of that one instance of repair or, more likely, through the repetitive use of the same type of repair in similar linguistic contexts (Skehan,

1998: 22). Though Skehan was fairly sceptical about the plausibility of these conditions being met within real L2 communication, I believe that these requirements place Swain's Output Hypothesis in the context of interaction which is possible in a language classroom and suggest the conditions which need to be incorporated into any grammar teaching framework, if it is to address the effectiveness of learners' cognitive processing for L2 acquisition.

In contrast to Skehan (1998), Gass and Mackey, among other researchers, actively support the potential of language development during interaction, which was first proposed by Long (1996). In his Interaction Hypothesis, he advocated the effectiveness of connecting L2 input with output for the development of learners' interlanguage, as well as of employing learners' cognitive capacities (see overview in Gass & Mackey, 2007, and Loewen & Sato, 2018). The most important part of the interaction, which is different from other forms of output, is considered to be "the negotiation of meaning", which Long (1996) defines as "the process in which, in an effort to communicate, learners and competent speakers provide and interpret signals of their own and their interlocutor's perceived comprehension, thus provoking adjustments to linguistic form, conversational structure, message content, or all three, until an acceptable level of understanding is achieved" (Long, 1996: 418). In other words, during the interaction, repairing a breakdown in communication, caused by the gap in learners' interlanguage, is a joint effort, and, thus, is potentially more effective.

Having analysed an abundance of examples from empirical studies, Gass (2015) concludes that interaction "pushes" learners to test their hypothesis and modify output (as in Swain, 1985, 1995) *more* than a one-way production, because it provides external feedback (Gass, 2015: 185). For that reason, she adds eliciting feedback as one more function of output; clearly focusing on external feedback (as in Swain & Lapkin, 1995, see 3.3.3). It is possible that external feedback provides more cues for an L2 learner to modify their output, quite likely in addition to the internal feedback, discussed in 3.3.3, thus potentially resulting in a higher rate of successful modifications. According to Gass, during interaction, learning happens "through the cycle of utterance, feedback, noticing of feedback, and modification" (Gass, 2015: 186). It is important to realize that 'noticing' here has a slightly different function, as external feedback normally implies the need for modification, unlike the internal, which sometimes could question the necessity of the change. Therefore, interaction increases the effectiveness of the 'hypothesis testing' (as in Swain, 1985) and probability of successful modification of learners' output. Consequently, it appears feasible that feedback could focus on grammatical form and be one more way of mobilizing learners' attentional resources and directing them towards, say, inflection, this way facilitating 'noticing', which in

the long run could lead to changes in learners' interlanguage. Moreover, feedback during the in-class interaction could potentially increase the opportunities for the outcome of modifications to become more salient and leave 'a trace' in learners' memory, as Skehan (1998) suggested (see above) as a condition for the effectiveness of acquisition, thus optimizing grammar information processing in a classroom environment. Gass (2015) states directly, that "the Interaction Approach relies on the efficacy of feedback" (Gass, 2015: 192). So, if the proposed teaching framework is to summon all cognitive resources available to learners for the most effective L2 acquisition, then external feedback on grammar forms should be an inseparable part of every activity.

Mackey (2006) investigated the relationship between feedback, 'noticing' and language development, during interaction. Her 28 participants, who were adult learners of English in US, of intermediate level, took part in specifically designed interactive activities, led by two teachers and the researcher. They were split into two groups – the experimental group received feedback on their use of Plural, Past Tense and question formation, while the control group did not. 'Noticing' was self-reported by learners, while the extent of language development was measured by pre-test and immediate post-test, which consisted of oral production tasks, designed to elicit the target forms. Mackey (2006) found that higher level of 'noticing' are undoubtedly associated with higher levels of learning; however the relationships between different components of interaction were quite complex. For example, there was a relationship between 'noticing' and feedback, as well as between 'noticing' and language development, but only for one of the grammatical features, which was question formation. Also, for the same feature (question formation), more 'noticing' was reported. Morphology appeared to be a problematic area, where extensive 'noticing' did not necessarily correlate with the increase in accuracy (it was strong and positive only for one participant). The explanation offered by Mackey (2006) was a higher salience of question forms due to syntactic movement, as well as to task demands, as questions were paramount for the interactive activity. One of the other inferences that could possibly be made, might be that 'noticing' is more effective when learners are able to concentrate on one particular form at a time. The other possibility is that inflection might take longer to develop than other types of functional morphology.

Furthermore, Gass and Mackey (2020) examine the development of the originally proposed Interaction Hypothesis into Interaction Approach, which they connect to learning. They follow Gass (2015) in that, during interaction, "the exposure to language, the production of language and the feedback on the production" are mediated by learners' cognitive mechanisms, eventually resulting in acquisition (Gass & Mackey; 2020: 194). However, the

researchers clearly state that interaction by itself might not always be effective for language learning and that Interaction Approach especially “values pushed or modified output, or that output which involves a learner attempting to go beyond his/her current level of knowledge” (Gass & Mackey, 2020: 212). They also stipulate that interactive events could include both explicit and implicit information, but do not specify what kind of knowledge results from it, though there are claims that both explicit and implicit knowledge benefit from feedback during interaction (Gass & Mackey, 2020: 213). Though the concept of learning is difficult to operationalize, a few studies indicate that there is a relationship between modified output and learning (see a summary in McDonough, 2005). Lightbown (2008) discusses the positive influence of interaction, as it generates multiple retrieval of language information and encourages deeper levels of processing through establishing multiple connections within neural networks.

How language development is impacted specifically by modified output, is analysed by McDonough (2005) through the prism of Levelt’s model (see 3.1), but in addition to de Bot (see 3.3.3), she specifies the mental processes that take place *after* ‘noticing’, in more detail. Assuming that the monitor, supervising overall speech production, consists of three components – the message, the grammatical and the articulatory, she sees three possible scenarios for modification of the output. First, the modification could be performed within the grammatical component *before* the articulation stage (as in Swain & Lapkin’s (1995) internal feedback). Second, the modification could take place in the same grammatical component *after* a gap or a fault is identified in the form that has already been produced, and the output is redirected back to the grammatical component for adjustment of functional issues. The third option, available to an L2 learner, could be when the produced output is sent back to the message component for re-formulation and then would go through the grammatical component again but, this time, as a new message (McDonough, 2005: 82). These three options present three schemes of retrieving the previously encoded grammar information with the aim of production, thus constituting different paths for recall. McDonough (2005) demonstrates that once a form has been produced, it is then produced repeatedly; this reinforcing the idea of multiple retrieval that strengthens the retrieval path on the way to proceduralization. This view of output modification offers a plausible explanation of how interaction could lead to changes in learner’s accuracy during language production and potentially register these changes in learners’ memory. Though McDonough does not split the grammatical component of Levelt’s speech production model into Function Assignment and Constituent Assembly, her discussion clearly points out that this is when grammar forms and functions have to connect and align with the message. This appears to be the part of

processing, which, as I argued in 3.2., is different from cognitive processing involved in the acquisition of other skills; and interaction seems to tap into it.

Furthermore, McDonough (2005) investigated the extent to which modified output, negative feedback or the combination of both result in language development. As a measure of learning, she used stages of development of question formation, testing 60 Thai learners of English, who took part in interactive tasks led by native speakers of English. The participants were split into four groups, which were administered four different treatments: 1) enhanced opportunity to modify (when the form was repeated back with the opportunity to correct the form); 2) opportunity to modify (with clarification requests and learners' follow-up output); 3) feedback *without* opportunity to modify (when the form was repeated with no opportunity for further output); 4) no feedback. Following the treatments, participants' oral production of target questions was tested individually, using tasks which differed from the treatment tasks. McDonough (2005) found that different learners used different paths for modifications (as discussed above) and produced modified output even with no feedback, but only clarification requests correlated positively with language development. (The scope of this study does not allow me to discuss different types of feedback in detail, which is a separate area of research - see a summary and the discussion in Gass & Mackey, 2020: 199-203). However, the only significant predictor of question development was modified output that involved the formulation of more advanced questions in response to negative feedback, this constituting empirical evidence for the potential effectiveness of interaction in a classroom.

Initially, the Interaction Hypothesis research was primarily looking at interactions between learners and native or more proficient L2 speakers. However, then the distinct shift from laboratory and naturalistic research toward real classroom settings occurred, producing a whole range of Instructional SLA studies that actively examined the effects of various interactive tasks on learners' L2s (see Pica, 2013, for a summary, and Philp, Adams & Iwashita, 2013, for an array of relevant studies). Philp, Walter and Basturkmen (2010) investigated specifically, how effective L2 peer interaction could be in promoting language development. They recorded 12 one-hour lessons of "French for Business" course over the period of three weeks, with 30 undergraduate students in the class, but only eight of them participating in the study. Within 26 interactions, 33 language-related episodes were registered; these are defined as instances when participants discuss the language that they produce, and when they correct themselves or their interlocutors (Swain & Lapkin, 1995: 326). Only one of those episodes was initiated by the teacher, half were instigated by the speakers and the other half by the listeners, that is, involved correcting others. This data, by itself, indicate that L2 learners are able to pay attention to linguistic issues without

participation of native speakers, teachers or more advanced learners. Moreover, Philp et al (2010) showed that only two of the total episodes did not result in successful repair, thus further demonstrating that peer interaction can be beneficial for addressing gaps in L2 and for resolving breakdowns in communication. In addition, the tasks that were performed by the students, were aimed at fluency and mastering of vocabulary. That is why, it is logical that the majority of language-related episodes were vocabulary-orientated and only seven (from the total of 33) were grammar-focused, mainly after teacher's reminder to focus on a particular form. This supports the earlier suggestions that explicit explanations and the encouragement of 'noticing' increase the probability of modifications of linguistic forms by L2 learners, thus reinforcing the importance of these elements for the proposed teaching framework. Furthermore, Philp et al (2010) argue that the episodes related to the discussion of form depend to a great extent on the task, illustrating this by past narrative instigating a number of modifications of past tense verbal forms. This indicates that interactive tasks need to be aligned with the grammar features which are introduced, however, these fine adjustments could not be investigated further within the scope of this study. Nevertheless, the general idea of task alignment is of extreme significance.

Furthermore, WM and attention are considered to play an important part in L2 learners' gains related to their interlanguage development during interaction. For example, Robinson (1995) stated that "individual differences in memory and attentional capacity both affect the extent of 'noticing', thereby directly influencing SLA" (Robinson, 1995: 283). De Bot (1996), for example, speculated that the demand on learners' cognitive resources increases dramatically during the formulation stage of L2 speech production (see Kormos, 2014) discussed in 3.1; and McDonough (2005) discussed above), particularly when a gap or mismatch repair is attempted. Mackey et al (2010) consider WM responsible for distributing learners' processing resources, such as selective attention, and directing these toward the L2 form. They agree with Payne & Whitney (2002) in that that, during interaction, learners' WM has to hold the representations of the input and, at the same time, recall L2 grammar information from their LTM, while they formulate their response (Mackey et al., 2010: 505). Learners also need to select which input is relevant and which information from the input needs to be encoded into the LTM (Mackey et al., 2010: 506). Drawing on previous research, Mackey et al (2010) summarize that there is a relationship between the effectiveness of output modifications, as well as what learners gain from interactive feedback, and their WM capacity. The focus of their study was the connection between the amount of modified output and the capacity of learners' WM. They hypothesized that learners with higher WM capacity would produce *more* output modifications in response to feedback during the interaction, than those with lower WM characteristics. Unlike in

McDonough (2005), the link between modified output and learners' language development was not investigated. Neither was the type of feedback; finally, all participants were provided with the opportunity to modify their output. Mackey et al (2010) speculated that, potentially, WM capacity might influence which path learners choose at the time of output modification (see discussion of McDonough (2005) above), while message is being sent for "reprocessing" (Mackey et al., 2010: 508). In order to provide feedback which is similar and as naturalistic as possible, in contrast to Philp et al (2010), Mackey et al (2010) used four native speakers to communicate with learners during their task-based interaction on one-to-one basis, rather than examining peer interaction in a group. For the experiment, 42 English-speaking learners of L2 Spanish were recruited from a US University, during their fourth semester. Four different tasks were used for treatments, but none of them aimed to elicit a particular form. All two-hour dyad interactions were recorded, transcribed and coded. The researchers managed to elicit over 500 instances of modified output, in half of the contexts where modifications were possible. The significant positive moderate relationship was found between WM scores and the amount of modified output that participants produced. Mackey et al (2010) further discuss that different types of feedback might put different demands on learners' cognitive resources. Furthermore, they discovered that different types of output (recast repetition and change of learner's own output) correlate with different aspects of WM, reserving more detailed investigations for future research.

The results of Mackey et al's study (2010) offer an explanation of different responses to feedback and might indicate that different learners might benefit from interaction to a different extent. Though this experiment points towards the link between WM and variability in L2 learners' learning, WM remains a "fixed trait" (as in Juffs & Harrington, 2011, discussed in 3.3.1.). Therefore, the only inference that is possible to make for instructional settings, is the necessity to focus learners' attention, during interaction, on particular forms that are being studied, with the aim of mobilizing their cognitive resources and help them develop a strategy to focus on form during any interaction by exercising 'noticing', which could potentially lead to larger gains for their language development.

The above discussion allows me to conclude that meaningful production of language within interaction activities create abundant opportunities for learners to recall the encoded grammar information (that is, declarative knowledge) in the classroom settings. They seem to learn most through modifying their output by "reprocessing" it, often in response to the external feedback, but sometimes via their own internal feedback. For that reason, interactive activities are to be an integral part of the proposed teaching framework, and abundant feedback needs to be provided by teachers, as well as to be encouraged among

the peers. However, as the response to feedback can vary quite considerably, it is essential to align interactive activities with particular grammar topics, and, more importantly, encourage learners to 'notice' the form that they study, in both the input and the output. Some kind of pre-planning could possibly form the expectation of the need to retrieve and produce the form in the output, similarly to how 'noticing' forms the expectation of the need to recognize the form in the input. Thus, explicit instruction, prior to interactive activities, should potentially increase the amount of modified output.

To sum this up, the combination of 'noticing' (as a kind of recognition) with interaction (instigating multiple recall) in classroom activities appears to provide an effective way of integrating focus-on-form with focus-on-meaning, as well as optimizing learners' processing of grammar features. It is very important that interaction appears to be able to address the issue of applying the encoded grammar rules to new contexts and assign the encoded suffixes to new lexical items during Function Assignment and Constituent Assembly stages (see Bock & Levelt's (2002) model in Figure 2 in 3.1.), which differs speech production from other cognitive skills. It is also possible to infer that interaction, which is organized to elicit particular grammar forms (e.g., inflection), could help L2 learners form representations in procedural memory, required for proceduralization. De Bot interpreted acquisition as the process "when new words are formed through the application of existing rules or the combination of morphemes previously acquired" (de Bot, 1996: 550). He further suggested that "these newly formed elements can move from a declarative phase to a procedural phase" (de Bot, 1996: 550).

3.4. Conclusion

Having discussed some relevant research in L2 grammar teaching, L2 grammar learning and SLA perspectives of grammar acquisition (see Chapter 2), I analysed some mental processes that are involved in acquiring L2 forms and possible ways of optimizing those in a language classroom (see Chapter 3). That allowed me to identify two models which I used to explain the processing of grammar information by L2 learners, namely WM model by Baddeley (1992) and the speech production model by Bock & Levelt (2002). In Section 3.1, I identified six stages, through which an L2 grammar form (e.g., inflection) has to go through to be produced by a learner – the first three are subprocesses of encoding and the last three are subprocesses of retrieval.

The cognitive research in psychology on memory suggests that the characteristics of WM are connected to the effectiveness of learning. However the relationship is complex and the

empirical studies seem to point out that WM is only one of many factors that affect language information processing, particularly for morphological inflections. Nevertheless, to increase the effectiveness of inflection production, both of the main processes involving WM, namely, encoding and retrieval need to be optimized. Also, the limited capacity view of WM adopted here, leads to two important inferences – first, the instruction needs to ensure that the learning material is processable for learners; and, second, instruction needs to consider that lexical items and grammar, if presented explicitly, as required in most beginner lesson materials, would compete for the same limited WM resources, thus new grammar needs to be initially taught using familiar vocabulary.

The cognitive view of acquisition, particularly Skill Acquisition Theory, provides some insights into how grammar information is stored in learners' memory, introducing the concept of proceduralization, which is assumed to lead to more automatized performance through faster retrievals of required information. It appears that my conclusion, made in 3.1, about the need of optimizing encoding and retrieval of grammatical features, is in line with the stages of proceduralization, defined by cognitive psychology. Thus, multiple attempts at performing an action following the explicit instruction (as seen by cognitive psychology), could be interpreted as multiple retrieval of previously encoded grammar information, with the aim of producing a certain grammar feature, say, inflection. However, I argue that the Skill Acquisition Theory does not account for two of the six stages of the L2 form production (see 3.1.), namely Function Assignment and Constituent Assembly, which deal specifically with case function and case form.

Therefore, to find out which kind of language activities would increase the effectiveness of encoding and the retrieval of case function and form, certain elements of SLA models are also adopted here. The Noticing Hypothesis (Schmidt, 1990) and some parts of Input Processing Theory (VanPatten, 2004) suggest that 'noticing' is an effective facilitator of recognition, while the Output Hypothesis and Interaction Approach appear to address the multiple recall. This brings me back to the discussion of the integration of teaching grammar with teaching speaking (see 2.1). It appears that Interactive Approach in SLA provides support to what researchers in language pedagogy have been moving towards - that is, the integration of teaching of form and meaning through oral production, particularly in communication.

Furthermore, some important findings in psychology and psycholinguistics point towards the necessity to address mechanisms of learners' processing, emphasizing the crucial role of strengthening the neural paths during multiple retrieval of encoded information (see 3.2 and

3.3.2), both recognition and recall. This appears to materialize in the integration of processing grammar features in both language input and learners' output, with 'noticing' being crucial for its effectiveness. Moreover, there are suggestions that multiple retrieval of the encoded declarative knowledge could help L2 learners form representations in procedural memory, required for proceduralization, leading to more automatized use of grammar features, which is the ultimate purpose of teaching grammar.

Below, I outline a number of important inferences that I have made from the discussions above, which will underpin my new proposed teaching framework, described in Chapter 4:

1) In order to ensure form-meaning mapping, essential for the acquisition of L2 grammar, the framework is designed to systematically integrate the teaching of grammar with the teaching of speaking and communication skills, thus incorporating both Focus-on-form and Focus-on-meaning (see 2.1).

2) The combination of both explicit and implicit modes of teaching and learning appears to have the potential to facilitate the acquisition of L2 grammar in a classroom (see 2.2.). That is why, both explicit instruction and the activities providing opportunities for creative communication, need to be built into the framework.

3) To optimize learners' processing of grammar features, instruction needs to first increase the effectiveness of the encoding of both the explicit grammar information and the language information that acts as an L2 input (see 3.3).

4) To address learners' WM limited capacity at the point of encoding (see 3.1), explicit information (explicit grammar rules) must be presented in smaller processable amounts but with smaller intervals between presentations.

5) Both types of retrieval are to be optimized – recognition and recall. To facilitate recognition of grammar forms, 'noticing' needs to be encouraged throughout the instruction (see 3.3).

6) To strengthen retrieval paths (which could potentially lead to proceduralization, discussed in 3.2), recall needs to be active (rather than passive) and multiple (see 3.3). The conditions for this are provided during meaningful interaction (see 3.3), which needs to become an integral part of every grammar topic. The interactive activities need to be aligned with explicit grammar presented, frequent feedback

needs to be provided and ‘noticing’, again, must be encouraged. Both Function Assignment and Constituent assembly have to be attended to (see Figure 2).

7) In order to enable learners to apply encoded grammar rules for morphological case and function to new lexical contexts, the proposed framework needs to assist learners in gradually building their interlanguage system, possibly through creating mental representations of L2 grammar features (see 2.3 and 3.3). Both ‘noticing’ activities and modification of output during interaction, are suggested to have the potential to contribute to that. In addition, linking specific inflection to broader grammar concept of case and the language system as a whole, in order to encourage deeper levels of processing, is to be part, to various extent, of all classroom activities.

Chapter 4

OPTIMIZING THE TEACHING OF L2 GRAMMAR, WITH PARTICULAR FOCUS ON RUSSIAN CASE INFLECTION

Before I outline the main aspects of my new proposed teaching framework, I will give a rather brief overview of the concept of case and Russian case system in particular, as it will be used to illustrate various aspects of the framework, as well as the basis for the experimental part of the current study. Parallel to this, I will examine how the issues of inflection acquisition and the processing of linguistic form, which were discussed above, are reflected in the learning of Russian cases. After that, I would put forward my new innovative teaching framework, describe its main principles and explain how it could make L2 learners' processing of grammar features, and Russian case inflection in particular, more effective.

4.1. Meanings/functions and forms of Russian cases and studies in the learning of Russian case inflection

4.1.1. The concept of case and some definitions

There are volumes dedicated to the explanations of the complexities of grammatical case systems, which have been actively investigated in linguistics for a number of decades, with several theories posited to explain the underlying principles of case assignment and case marking (Babby, 1980; Baker & Vinokurova, 2010; Chomsky, 2001, 2004; Jakobson, 1984 [1958]; Pereltsvaig et al., 2018; Pesetsky, 2013; among others). However, as Malchukov and Spencer (2009) stated, "the concept of case and its relation to grammatical relations, meaning, and morphological form remains elusive and controversial" (Malchukov & Spencer, 2009: 1). The definitions of case are as varied as its typology and range from "(inflected) form of a nominal word" and "property of a noun phrase (determiner phrase)" to "[c]ases are usually suffixes" (Spencer, 2009: 185). Haspelmath (2009) gives a comprehensive analyses of terminology relating to case, as well as a general definition, widely accepted by linguists, – "inflectional category system expressing dependency relations" as well as "the individual inflectional categories or values of that system" (Haspelmath, 2009: 505). Though this definition fits a huge variety of case systems, it is rather general for the purpose of the present study, which focuses on how learners process and acquire case in a classroom. More detailed descriptions and typological classifications of cases, to a large extent, depend on a view of language acquisition, discussed in 2.3. In generative grammar, the assignment of case is normally explained with the help of thematic roles closely related to syntactic structures (e.g., Abraham & Leiss, 2013), while functionalists define cases in terms of

meanings and semantic connections (e.g., Janda, 2002, 2013; Kempe & MacWhyney, 1998).

Despite the abundance of research on case, there does not appear to be many studies researching how the findings of the above investigations could benefit learners of Russian, with cases remaining one of the most challenging aspects in Russian language learning (e.g., Arnett & Lysinger, 2013: 135; also see Gor, Chrabaszcz & Cook, 2019; Izurin, 2013; Janda, 2002; Rifkin, 2005). As this study focuses on optimizing learners' processing of grammatical form, I will examine the complex concept of case from the point of learners' processing and their acquisition needs. In order to do that, I will attempt to simplify it by reducing the enormous amount of information, accumulated about case, to two basic aspects, namely, the *meaning* (or the function, or the theta-role) and the *form* (or marking, or suffix, or ending). From the discussion of Bock & Levelt's speech production model in 3.1., it is apparent that these two concepts affect any learner's decision every time s/he uses a case, and, thus, need to be first encoded and then repetitively retrieved. DeKeyser (2005) confirmed this by saying that three main factors determining the difficulty of grammar – "complexity of form, complexity of meaning, and complexity of the form-meaning relationship", that is in addition to "grasping form-meaning relationship" (DeKeyser, 2005: 3).

Further in this section, I will briefly review how these two aspects are presented in grammar references, and how various aspects of that information would affect learner's encoding. Then, I will examine what information on case is provided in course books and what kind of retrieval it stimulates, in order to see which aspects of instruction need to be addressed within the proposed framework. Finally, I will describe some relevant studies that investigated the processing of Russian case inflection, to see how they could contribute to the discussion in the previous two chapters.

4.1.2. Case meanings/functions and forms in Russian and learners' processing

In this section, I will introduce some basic information about the six Russian cases commonly acknowledged in Modern Russian, namely, Nominative, Accusative (Acc.), Prepositional (Prep.), Instrumental, Genitive and Dative (e.g., Haspelmath, 2009). The present study will follow the widely-accepted terminology, summarized by Haspelmath (2009), though the case names in some of the figures from other volumes might differ occasionally (e.g., Prepositional and Locative could be differentiated).

With regard to meaning, cases are given thorough explicit treatment in most traditional Russian grammar descriptions, with fairly comprehensive lists of nuanced meanings accompanied by a generous supply of authentic examples (e.g., Offord, 1996; Shvedova, 1989; Timberlake, 2004; Wade, 2011) (also see Durst-Andersen & Lorentzen, 2017, for a review). This level of detail in descriptions, abundant linguistic terminology and complexity of structures of exemplar sentences are far beyond the beginners' level, when the cases are introduced and are expected to be proceduralized, making this kind of references inappropriate for low-level learners. Some of the introductions of the editions above, as well as their academic reviews, clearly state that these are for students of "post-introductory stage" and "a reference aid for teachers, translators and interpreters" (e.g., Ward, 1993: 534). Also, the case forms from authentic examples from classic Russian literature would be unlikely to be processed and retained by, say, ab-initio language students, as their cognitive resources would be used up by decoding numerous unfamiliar vocabulary and possibly complex discourse.

An example of the functionalist description of Russian cases is Janda and Clancy's (2002) The Case Book for Russian, where case meanings appear more usage-orientated and learner-friendly. In addition, the explanations are accompanied by pictograms, schematizing the function of each case in context, assisting their comprehension. Finally, a basic meaning is singled out from various submeanings, which could be helpful at initial stages of learning. Despite all the enumerated advantages and considerable number of grammar exercises, Janda and Clancy clearly state that the book does not have a purpose of helping learners to acquire Russian case suffixes, aiming at enhancing their understanding of the holistic concept of case and its application in Russian, thus concentrating on the explicit part of learning. The list of the basic case meanings for Russian, offered by Janda (2002), is presented in Table 1 - I will discuss this further in this chapter for the purpose of adapting it for my proposed teaching framework.

Table 1

Core Meanings (Functions) for Russian Cases (Janda, 2002: 2)

Nominative:	**a name** (naming, subject); **an identity** (predicate nominative**)**
Genitive:	**a source** (withdrawal); **a goal** (approach); **a whole** (possession/'of', quantification); **a reference** (lack, comparison, near)
Dative:	**a receiver** (indirect object); **an experiencer** (benefit, harm, and modal uses); **a competitor** (equality, submission, domination)
Accusative:	**a destination** (movement, direct object, points in time, durations, distances, amounts)

Locative:	**a place**
Instrumental:	**a means** (means, instrument, path, agent); **a label** (predicate instrumental); **an adjunct** (preposition s 'with'); **a landmark** (prepositions of proximal location)

As suggested in 3.3. (see 3.4. for a summary), creating numerous opportunities for encoding and retrieving of these meanings/functions, in terms of Function Assignment, at least, needs to be an important part of beginners' grammar instruction for case. L2 learners can practice the Function Assignment during their oral production (see Figure 2 in 3.1.), gradually adding more and more case meanings/functions, to their mental database.

(To avoid confusion between the two interpretations of the term "meaning", which appear to emerge here, from here onwards, I would refer to "meaning" in the sense of lexical meaning, associated with semantics, as used in Chapter 2 and Chapter 3; while case meanings, as in Janda (2002), would be called "case functions".)

In addition to case functions listed in Table 1, each Russian case is used in a number of pragmatic contexts, which do not appear to directly relate to these main functions. For example, Russian Prep. case, as well as having the main function of "a place" (or location), is used for some time references (e.g., months and years), for transport, for events and for musical instruments (after the verb "to play"), among other contexts. Similarly, Acc. is used for days of the week and games. These contexts considerably diversify case function information, complicating the Function Assignment process for L2 learners.

Moreover, it is possible that Function Assignment could be more challenging in some case contexts than others, as some case contexts appear to involve nouns belonging to different types of lexical sets (syntagmatic sets, as in Jezek & Hanks, 2010). The difference is as detailed below.

- Closed lexical sets – that is, those that do not accept new members, thus, once studied, only include familiar items, e.g., days of the week.
- Limited lexical sets – that is, those that are larger and, though potentially open-ended, have well-defined list of frequently used items, many of which are familiar to learners, e.g., musical instruments.
- Open lexical sets – that is, those that have no limit on a number of lexical members and can include endless number of items, easily accepting new words, e.g., place names.

From the point of view of processing difficulty, these sets require different amount of cognitive effort. Closed sets, for instance, would be at the bottom of the scale, as the small number of items in a set (e.g., seven for days of the week) can allow chunking and their production might be less demanding than application of inflection rules to new items (see ACT-R processing model by Anderson & Lebiere, 1998, discussed in 3.2). On the contrary, open sets include unfamiliar items and both Function Assignment and Constituent Assembly would have to be executed for those items for the first time, thus requiring larger amount of cognitive resources than the other sets. Items from limited lexical sets can be produced as chunks with some potential for unfamiliar words. That is why case context might appear one of the factors that affect success of inflection production and investigating these might shed light on learners' challenges, thus possibly suggesting ways of improving their inflection accuracy.

Also, I decided to use the term "case context", rather than "case function" (as in 4.1.1), as the latter is based on syntax and is quite broad, as well as sometimes being difficult to comprehend for learners. For example, "a distination" for Acc. (see Table 1) can be easily confused by beginner learners with no linguistic background, with "a place" which is listed as a Prep. case function. On the contrary, "case contexts" are linked to vocabulary which require a particular case and its exact use, thus are connected to semantic meanings and can be easily distinguished. However, it is not the same as "semantic context", which might not necessarily be linked to a case, as, for example, both months and days of the week belong to the same semantic context of temporal references, while months require Prep. and days of the week need to go into Acc.

With regard to form, three main noun classes, or declension types, are traditionally recognized in Russian. Each of those has a separate set of suffixes for the six cases (often a fourth class and some subclasses are distinguished (e.g., Cubberley, 2002: 110-120)). Nouns are allocated to these classes, according to their Nominative suffixes, which, to a degree, are associated with one of the three genders – masculine (normally with no overt inflection, or zero suffix), feminine (normally marked by "a" or "â") and neuter (normally marked by "o" or "e"). This alignment though, is not straightforward, for example, Class II is represented by both masculine and neuter nouns; feminine nouns can belong to one of two classes (those marked by a/â - to Class I; those with no phonologically overt inflection - to Class III); but Class I, in addition to feminine items (marked by a/â), has a cluster of male persons, also ending in a/â. (Note that labelling of these classes could vary.) Table 2 presents basic categories of nouns included in each class.

Table 2

Russian Noun Classes (Declension Types), Based on Brooks & Kempe, 2008: 707

Gender	Feminine		Masculine		Neuter
Noun Class (declension type)	Class I	Class III	Class I	Class II	Class II
Singular Nominative	a / â	ø	a / â	ø	o / e

In addition, feminine nouns ending in “iâ” and neuter nouns ending in “ie” form distinctive subclasses, with two more sets of suffixes, different from standard paradigms for these genders.

From the discussion above, it becomes apparent that gender assignment in Russian is not always transparent, with the form of gender marking not providing a reliable cue for processing the function. This consequently affects the transparency of noun classes (declension types) and is discussed further in 4.1.2. Furthermore, Russian nominal inflection fuses the categories of gender and case - that is, one suffix incorporates both of these, as well as the number category, for example, “a” in Table 2, in the majority of cases, would indicate feminine, Singular, Nominative. (This is in contrast to agglutinative morphology, when each of those categories would be marked by a separate inflection, all sequenced one after the other (e.g., M. Martin, 1995 for the analysis of Finnish morphology).) In Russian grammar references, case paradigms are traditionally provided in list form for each of the noun classes, as in the example illustrated in Figure 3 taken from Cubberley, 2002: 112-113.

Figure 3

Example of a paradigmatic list layout for Genders/Noun Classes from a grammar reference

	Feminine (and Masculine and Common personal)
Singular	
Nominative	a
Accusative	u
Genitive	i
Locative	e
Dative	e
Instrumental	oj
Plural	
Nominative	i
Accusative	In. i; An. ϕ
Genitive	ϕ
Locative	ax
Dative	am
Instrumental	ami

	Masculine	**Neuter**
Singular		
Nominative	ø	o
Accusative	In. ø; An. a	o
Genitive	a	a
Locative	e	e
Dative	u	u
Instrumental	om	om
Plural		
Nominative	i	a
Accusative	In. i; An. ov	a
Genitive	ov	ϕ
Locative	ax	ax
Dative	am	am
Instrumental	am′i	am′i

Now I will attempt to analyse the Russian case system from the perspective of an L2 learner and their processing. In the present study, I will use the term “oblique case” for non-Nominative cases, as defined by Haspelmath (2009), though I am aware that there are other interpretations of the term. From the point of view of form, a simple arithmetical calculation (five oblique cases multiplied by three noun classes (without any subclasses) in Singular) would produce 15 form-meaning pairs. However, as Figure 3 illustrates, there is a fair amount of syncretism within the Russian case system, when the same inflection (overt or not) is used for different cases. For example, “e” marks both masculine Locative (that is, Prep.) and feminine Dative; “u” can indicate feminine Acc., as well as masculine Dative; and Class II masculine nouns have no overt inflection for either Nominative or Acc.. In relation to Russian cases, Janda (1993) noted that "no single case has a unique surface representation by means of which it is signalled.” (Janda, 1993: 10). It is obvious that, in the Russian case system, there is no one-to-one relationship between a case form and its function, and processing this information is likely to be very challenging for a beginner learner, who would need to identify the right context during the online processing.

Furthermore, Acc. has different inflection for animate and inanimate masculine nouns of Class II (see Figure 3), which brings one more category to be accommodated in learners’ LTM. In addition, monosyllabic masculine nouns (of Class II) distinguish between Locative (denoting location) and Prep. (used after the preposition “o” for “about”), adding to the amount of information which needs to make it through learners’ WM and be allocated within their LTM, contributing to the challenges of selecting a form during production. In addition, prepositions do not always indicate a particular case, as the same preposition can be used for different case, for example, preposition *v* can mean “in” and require Prep. or it can stand for “to” and need Acc.

To add to the above, Russian nouns have a separate paradigm in Plural, with many case forms overlapping for different genders, but others being different for different noun classes. Due to the limited scope of this study, the category of number will not be discussed any further and Plural case forms in Figure 3 are given for information purposes only.

To complicate things further, some nominal case suffixes in Russian can be the same as in adjective declensions, for example, *om* is used to make Instrumental for masculine nouns (Class II) (*s limon**om*** meaning “with lemon”), as well as making Prep. for masculine adjectives (*v bol'š**om** dome* meaning “in a big house”). This is not ambiguous for a native speaker, but could be very confusing for, say, an English speaking learner, as grammatical categories of noun and adjective are not as strictly defined in English and have no overt marker. Finally, if these diverse case forms are combined with the list of case functions (see Table 1), it becomes clear why Russian cases remain “the biggest obstacle faced by English-speaking students trying to learn Russian” (Janda & Clancy, 2002: 3), and similarly challenging for other L1s. It is easy to see that these numerous forms and functions of cases can get lost in any of the multiple stages on the way from being registered within WM, being encoded into LTM, being recalled from LTM with the aim of production, during the grammar stage of the actual production and, finally, during the articulation of the utterance (see 3.1.)

Moreover, phonology and orthography play a significant role in case inflection production, adding to the opacity of form-meaning relationships. The first and the most important phenomena is palatalization, which is one of the major parts of Russian phonological system, when the majority of Russian consonants (15) have two versions – hard (non-palatalized) and soft (palatalized), that is, pronounced closer to the palate. This manifests itself in Russian orthography in general, and that of inflection in particular. For example, in Table 2, both masculine Class II nouns and feminine Class III nouns have no overt inflection (zero suffix - ø). In reality, feminine Class III nouns always have a soft (palatalized) consonant at the end and have to have the Soft Sign in spelling, while masculine Class II nouns tend to end in a hard (non-palatalized) consonant, which does not require the Soft Sign. However, some of Class II masculines *can* have a soft (palatalized) consonant at the end, which results in two spelling versions for half of the masculine case forms. For example, the Genitive “a” inflection (see Figure 3 above) can only be added after hard (non-palatalized) consonants (e.g., the Genitive of *hleb* (for “bread”) is *hleb**a***, spelt as “хлеб**а**”), while the soft (palatalized) stems need to be followed by “â”, which is spelt as “я” (e.g., the Genitive of *aprel'* (for “April”) is *aprel**â***, spelt as “апрел**я**”) (see Table 3). In the same manner, Instrumental has “em” as a soft (palatalized) alternative to “om”, and Dative has two

Class II allomorphs of “u” and “û”, with two corresponding spellings. The same spelling versions apply to neuter nouns (Class II) with “o” and “e” marking, corresponding to hard and soft stems respectively. Palatalization has a similar effect on feminine Class I nouns marked by “a” or “â”, as “a” follows hard (non-palatalized) consonants and “â” is found after soft (palatalized) stems. Consequently, in Acc., Class I feminines can have “u” or “û”; in Genitive, “i” or “y”; and, in Instrumental, “oj” or “ej”. In addition, there are two more spelling rules (related to some fricative consonants, which also affect Genitive inflection of feminine nouns and Instrumental of all genders. As L2 learners normally process visual explicit information, orthography would be of extreme importance for case inflection production, as different allophones would participate in form-meaning mappings, thus, would have to be considered by a learner while selecting a form during the Constituent Assembly stage of production. Moreover, if the concept of palatalization is not established in the learner’s interlanguage, allophones might have some small potential of erroneously competing for different cases.

Table 3

Case Declension Paradigms for the Three Noun Classes/Declension Types, Including Allomorphs (Based on Brooks & Kempe, 2008: 707)

	Feminine		Masculine	Neuter
	Declension I (transparent)	Declension III (nontransparent)	Declension II	
Nominative	-a (-â)	ø	ø	-o (-e)
Genitive	-i (-y)	-i	-a (-â)	-a (-â)
Dative	-e	-i	-u (-û)	-u (-û)
Accusative	-u (-û)	ø	ø (inanimate) -a (-â) (animate)	-o (-e)
Instrumental	-oj (-ej)	-û	-om (-em)	-om (-em)
Locative	-e	-i	-e	-e

Another example of the effects of phonology on case inflection is the stress shift, which would require some cognitive resources during the articulation of case forms. For example, some monosyllabic Class II masculine nouns tend to shift their stress to the overt inflection in oblique cases, e.g., the Prep. form of *stol* (for “a table” or “a desk”) is *na stolé* (“on the table/desk"). However, there is no exact rule of which monosyllabic masculines are to shift their stress, and some nouns of this subclass do not perform the shift or shift it in some case

forms but not the others. Similarly, some double-syllable feminine nouns with the stress at the end (mainly of the Old Slavic origin), tend to shift their stress to the first syllable in Acc., e.g., “sredá” (for “Wednesday”) becomes “srédu”. The attention that has to be paid to stress in the production of Russian case forms is to be accounted for, when processing of these forms is considered.

Finally, I believe that gender of the inflected noun would also affect the Constituent Assembly stage during inflection production, as, for example, it would require an additional step of taking a feminine base-form marker (“a” or “â” for Class I feminine nouns) off before adding a case marker. For example, *sumka* (for “a/the bag”) would have a two-step inflection process: *sumk**a***[*fem.base form*] + “e” > 1) *v sumk_* + “e” > 2) *v sumk**e***[*Prep.fem*] (for “in the bag”). This step is missing during masculine case form production, as the majority of masculine nouns (Class II) have no overt marker (that is, have a zero-suffix). For example, *stul* (for “a/the chair”) has a one-step process, when the case suffix is added to the end of the base-form: *stul* + “e” > *na stul**e***[*Prep.masc*] (for “on the chair”). Therefore feminine case forms would require additional processing on behalf of an L2 learner. Thus, gender could be one of the factors affecting Russian case inflection production and will be investigated in the present study.

The above discussion further illustrates the complexity of Russian case inflection and the diversity of functions, linguistic categories and forms that need to be encoded and retrieved by L2 learners during the Function Assignment and the Constituent Assembly stages of oral production, as well as during the actual articulation (see Figure 2 in 3.1.). This further supports my suggestion that addressing learners’ processing load in teaching instruction has the potential to increase the accuracy of case inflection production. In addition, three factors have been identified which could possibly affect case accuracy rates, namely, gender, case contexts and familiarity of lexis. These will be investigated in the current study (see 6.7. and 6.8).

4.1.3. Case inflection encoding and retrieval in beginners’ Russian textbooks

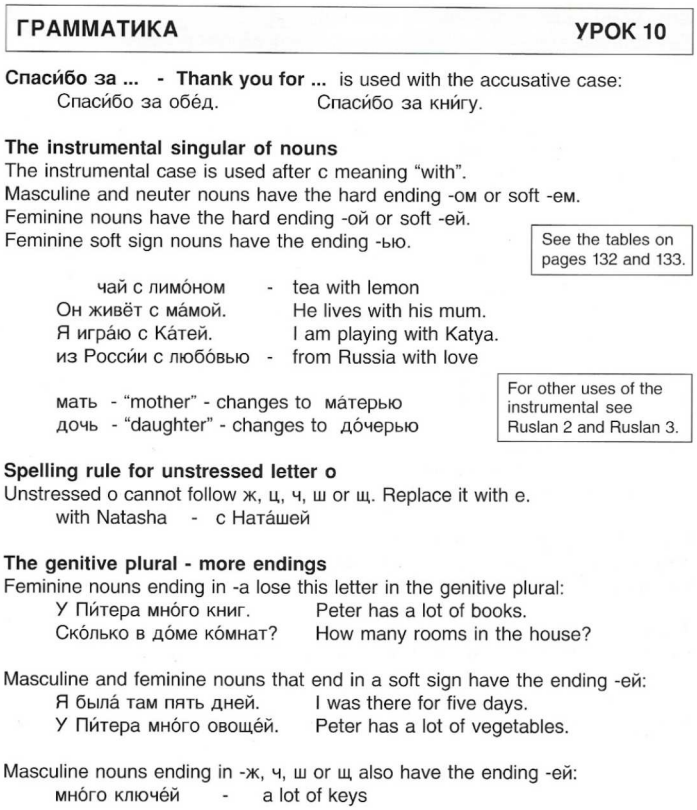
As the present study focuses on investigating the case acquisition at the onset of learning, I will examine only the textbooks, or course books, aimed at beginner learners, in order to see what grammar practice is provided for learning cases and how it addresses the encoding and multiple retrieval of case inflection, and which stage of case form production in speech they attend to.

The books used for standard Russian beginner courses at the English-speaking Universities, appear to primarily focus on the element of form, that is, on case names and sets of corresponding case suffixes, rather than on a concept of case (see Arnett & Lysinger, 2013, for a review of US published textbooks). Core (or nuanced) meanings, at least at the beginners' stage, are not explained as such, and usage is traditionally tied up to prepositions or certain words or phrases, requiring a particular case. For example, *s* (for "with") is often associated with Instrumental (though it can also stand for "from", when it would need Genitive); or *U menya net* (for "I don't have" or "I haven't got") or *mnogo* (for "many", "much" or "a lot of") are followed by Genitive (though countable and uncountable nouns would be inflected differently).

Traditionally, explicit grammar explanations are provided before practice exercises. Unlike grammar references, beginner textbooks present a set of case suffixes for a particular case within one lesson/unit. Though Singular and Plural case forms are normally introduced separately, sometimes adjectival case forms are introduced at the same time as nominal case forms (e.g., Robin at al, 2007). The explicit information in these books appears to be quite condensed for learners' processing, with the whole sets of suffixes for all three noun classes, including subgroups and allomorphs, presented in one go. See an example of a grammar section for one lesson in Figure 4.

Figure 4

Extract from a Grammar Section for One Lesson (Lesson 10) from Langran & Veshneva, 2008; p.120, a Course Book for Russian, Used by One of UK Universities.



ГРАММАТИКА **УРОК 10**

Спаси́бо за ... - Thank you for ... is used with the accusative case:
Спаси́бо за обе́д. Спаси́бо за кни́гу.

The instrumental singular of nouns
The instrumental case is used after с meaning "with".
Masculine and neuter nouns have the hard ending -ом or soft -ем.
Feminine nouns have the hard ending -ой or soft -ей.
Feminine soft sign nouns have the ending -ью.

See the tables on pages 132 and 133.

чай с лимо́ном - tea with lemon
Он живёт с ма́мой. He lives with his mum.
Я игра́ю с Ка́тей. I am playing with Katya.
из Росси́и с любо́вью - from Russia with love

мать - "mother" - changes to ма́терью
дочь - "daughter" - changes to до́черью

For other uses of the instrumental see Ruslan 2 and Ruslan 3.

Spelling rule for unstressed letter o
Unstressed o cannot follow ж, ц, ч, ш or щ. Replace it with e.
with Natasha - с Ната́шей

The genitive plural - more endings
Feminine nouns ending in -a lose this letter in the genitive plural:
У Пи́тера мно́го книг. Peter has a lot of books.
Ско́лько в до́ме ко́мнат? How many rooms in the house?

Masculine and feminine nouns that end in a soft sign have the ending -ей:
Я была́ там пять дней. I was there for five days.
У Пи́тера мно́го овоще́й. Peter has a lot of vegetables.

Masculine nouns ending in -ж, ч, ш or щ also have the ending -ей:
мно́го ключе́й - a lot of keys

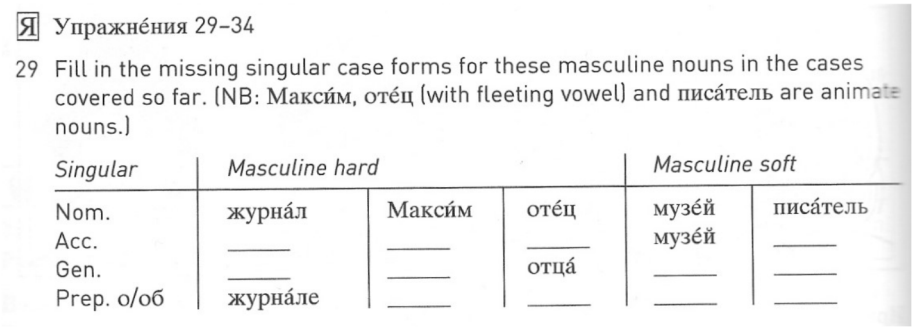
With regard to practice (see discussion in 2.1. and 3.3.3), according to my own small evaluation (see the list of evaluated textbooks in Appendix A), the majority of books offer traditional grammar drills, translations and PPP type practice (see 2.1.) (also, see Comer & deBenedette, 2011, for a short discussion, and Jackson & Kaplan, 2001, for an overview of classroom practices). In teaching materials published in Russia, the term "governing" of cases is often used to denote connections between a verb and its arguments. Consequently, cases are often practiced in verb-noun phrases, rather than isolated forms, which could potentially strengthen these connections, though grammar drills still seem to be a common type of practice (e.g., Zeng, 2017; Ryzhova, 2008). Overall, the majority of books contain a mix of different types of exercises and activities – from those deemed to be ineffective by research (see 2.1. and 3.1.), to those which are edging towards some kind of interaction, which was identified as the most effective (see 3.3.3.). I will illustrate with a few examples what kind of processing would be involved when learners carry out some typical activities which I found in the evaluated textbooks.

Some books still include declension exercises (see example in Figure 5), which represent passive retrieval of explicit information, based on memorization. As it is not encoded within any kind of context which would assist in retaining or retrieving it, this kind of activity is not associated with acquisition (see discussions in 2.1. and 3.3.1). Though grammar translation

method has been long discarded as a grammar teaching option, translation tasks are still present in textbooks. According to research, translators and proficient L2 speakers use different areas of the brain, which means that translation exercises train the skill which is not required for accurate L2 production.

Figure 5

Grammar Drill based on Noun Declension (Aizlewood, 2007: 122)



Я Упражне́ния 29–34

29 Fill in the missing singular case forms for these masculine nouns in the cases covered so far. (NB: Макси́м, оте́ц (with fleeting vowel) and писа́тель are animate nouns.)

Singular	*Masculine hard*			*Masculine soft*	
Nom.	журна́л	Макси́м	оте́ц	музе́й	писа́тель
Acc.	_____	_____	_____	музе́й	_____
Gen.	_____	_____	отца́	_____	_____
Prep. о/об	журна́ле	_____	_____	_____	_____

Traditional open-the-brackets exercises are often employed for grammar practicing (see example in Figure 6). Though these put case forms into sentences, thus giving them some context and linking them to case functions, they appeal solely to learners' declarative knowledge and, therefore, unable to contribute to the use of case forms in any of the four language skills, and consequently, would not be able to initiate "internalization" (R.Ellis, 2006) or proceduralization (see 3.2.).

Figure 6

An example of open-the-brackets exercise for the Instrumental case (Smith & Crosbi, 2002: 147; used in one of UK Universities)

Упражнение 7. Coordinated subjects. Кто с кем что делал? *Put the words in brackets into the correct form.*

(а) В субботу я был в театре с (брат) . . .
(б) Вчера я ужинал с (Борис) . . .
(в) Я говорил по телефону с (товарищ) . . .
(г) Мальчики ходили в музей с (учитель) . . .
(д) Вчера он был на стадионе с (отец) . . .
(е) Я хочу поговорить с (Катя) . . .
(ж) Отец ходил в цирк с (Ирина) . . .
(з) Он разговаривал с (профессор) . . .
(и) Мать гуляла в парке с (сын) . . .

Exercise 7. Coordinated subjects. Who did what with whom?

(а) On Saturday I was in the theatre with (brother)
(б) Yesterday I had supper with (Boris)
(в) I spoke on the phone with my (mate)
(г) The boys went to the museum with (teacher)
(д) Yesterday he was at the stadium with (my father)
(е) I want to speak with/to (Katya)
(ж) My father went to the circus with (Irina)
(з) He spoke with (professor)
(и) The mother had a walk in the park with (her son)

In more recent publications, there appear exercises that encourage ‘noticing’ of case forms during recognition (see example in Figure 7). If analysed from the processing point of view though, the meaning of these sentences could be extracted just from the lexical units used, even if the inflection was not present, making it redundant, which is suggested as being not effective for encoding (VanPatten, 2004). In addition, the input consists of isolated sentences not connected to each other by some common theme or plot, as well as their content being of no interest or relevance to learners, thus, might not leave a strong “trace” (as in Skehan, 1998; see 3.3.2) in their LTM. For this kind of activities to contribute to effective recognition of case forms and possibly forming their declarative representations required for acquisition (see 3.2), they would need to incorporate case forms into a wider context on the text level, this way providing more meaningful connections. Also, as it was discussed in 3.3.3., they should be closely followed by production activities, in order to contribute to forming procedural memory.

Figure 7

Grammar Exercise Consisting of Isolated Sentences, for Practicing Dative Case (Kudyma, Miller & Kagan, 2016: 174)

13-5. Кому́? Read the following sentences and circle the nouns in the Dative case. Explain the endings. Underline the verbs after which the Dative case is used. See Grammar comment 13-1.

1. Бра́ту 25 лет.
2. Макси́м рассказа́л дру́гу об Аме́рике.
3. Мы не говори́ли дя́де о но́вой кварти́ре.
4. Он не сказа́л нам, как его́ зову́т.
5. Ольге не нра́вятся мои́ друзья́.
6. Вы да́ли Петру́ Петро́вичу ваш дома́шний телефо́н?
7. Оте́ц дал сы́ну де́ньги.
8. Ма́ма купи́ла Ири́не но́вый компью́тер.

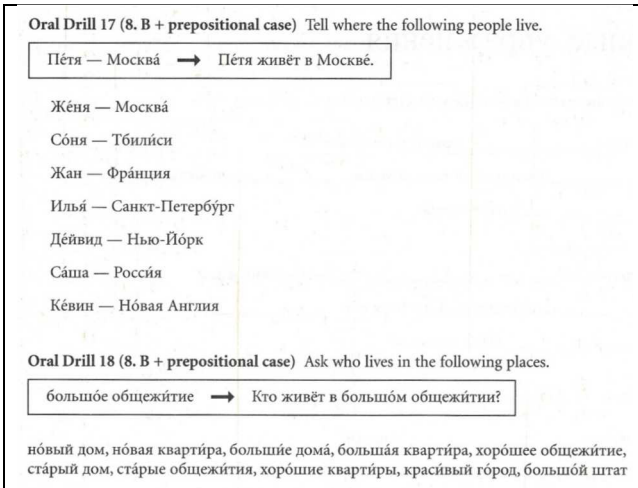
1. My brother is 25.
2. Maxim told a friend about America.
3. We did not tell our uncle about the new flat.
4. He did not tell us, what his name was.
5. Olga does not like my friends.
6. Have you given Pyotr Petrovich your home number?
7. The father gave his son some money.
8 Mum bought a new computer for Irina.

Some textbooks have started introducing structured oral production activities for grammar practice, as speaking activities were traditionally meaning-orientated. These are mainly of the PPP type (see 2.1.) (see example in Figure 8), but are a step forward, as they involve form-focused oral production, without repeating a ready-made phrase; thus, encouraging some cognitive processing. During these activities, learners have to retrieve a correct case suffix, appropriate for a noun class/gender. Thus, they perform Constituent Assembly, while case function is already selected for them. This might be a good step-up before interaction, but for Function Assignment to be activated, there need to be at least two case functions for

learners to choose from, therefore two different case contexts included into one activity would be more effective. As lexical items are already provided, the vocabulary recall is not competing with the recall of the form. On one hand, this way, leaners' cognitive resources are fully allocated to form, as there is no lexical retrieval to perform. On the other hand, the mechanical nature of the output and absence of meaning-orientated relevance do not allow these to be classed as task-based production (see 3.3.3), but still might be used as a step-up activity, before the interaction.

Figure 8.

Oral Production Activity for Practicing Prep. Case (Robin et al., 2007: 65)



Oral Drill 17 (8. В + prepositional case) Tell where the following people live.

Пётя — Москва́ → Пётя живёт в Москве́.

Же́ня — Москва́
Со́ня — Тбили́си
Жан — Фра́нция
Илья́ — Санкт-Петербу́рг
Де́йвид — Нью-Йо́рк
Са́ша — Росси́я
Ке́вин — Но́вая Англия

Oral Drill 18 (8. В + prepositional case) Ask who lives in the following places.

большо́е общежи́тие → Кто живёт в большо́м общежи́тии?

но́вый дом, но́вая кварти́ра, больши́е дома́, больша́я кварти́ра, хоро́шее общежи́тие, ста́рый дом, ста́рые общежи́тия, хоро́шие кварти́ры, краси́вый го́род, большо́й штат

Oral Drill 17.
Petya - Moscow → Petya lives in Moscow*[Prep.fem]*

Zhenya – Moscow
Sonya – Tbilisi
Jean – France
Ilya – St. Petersburg
David – New-York
Sasha – Russia
Kevin – New England

Oral Drill 18.
Big hostel → Who lives in big hostel?

new hose, new flat, new houses, new hostel, old house, old hostels, new flats, beautiful city, large state

Furthermore, Kudyma et al (2016) offer learners to conduct surveys among their peers at the end of each unit and report on it back to the class. This activity has the characteristics of interaction, discussed in 3.3.3, as long as the "surveys" are conducted in the form of interviews (that is, orally) and learners modify their output to repair gaps in communication, rather than adhering to English. Finally, it would be more effective if tasks could be varied, rather than being identical in each unit, though the content of the "surveys"/interviews would still be different and relevant to learners, as well as being aligned with the grammar studied.

Overall, my small evaluation of grammar information and activities in beginners' textbooks shows that explicit grammar explanations do not normally address leaners' processing restrictions, such as WM limited capacity (see 3.1), and mainly rely on memorization. As for grammar practice, it appears that many traditional exercises in the evaluated materials are aimed at maintaining declarative knowledge. Some exercises that initiate 'noticing', were found, but their effectiveness could be improved. There are examples of form-orientated oral

production, which helps learners perform Constituent Assembly (as in 3.1) but no activities specific for training Function Assignment were identified within the scope of the evaluation performed. Only one textbook suggested form-orientated task-based peer interaction, which was replicated in each unit. This means that the activities based on multiple retrievals (whether recognition or recall), which are considered most effective for acquisition (see 3.3.3), rarely present, if at all. In 4.2, I will propose some possible ways of addressing this deficit. I believe, that the proposed teaching framework needs to incorporate a range of activities, addressing different stages of processing (see 3.4.) to prepare learners for going through all of them during the interactive activities, which would need to be an essential part of each session.

4.1.4. Some studies investigating the learning of Russian case inflection

In recent years, a few studies appeared, investigating specifically L2 learners' performance on Russian case inflection from the point of view of learners' processing. Here, I will analyse a few which are relevant to the current study.

In a series of psycholinguistic experiments, Kempe and Brooks (2008) investigated English-speaking adult learners' performance with regard to generalization of inflection patterns in miniature language, which was subpart of Russian case system. Their two groups were learning transparent and non-transparent case forms, both in laboratory conditions, expected to deduct the inflection rules from the input, with no explicit instruction. No significant difference was found between the groups, leading to the conclusion that the majority of adult learners struggle to extract complex inflection patterns from the input, even if the input is structured (thus, battling with forming abstract declarative knowledge (see discussion of Kirkhart's study (2001) in 3.1), or possibly mental representations). This conclusion reassured me in my decision about making explicit instruction an inseparable part of my proposed teaching framework (also, see 2.1. and 3.2. for discussion).

Brooks & Kempe's (2013) subsequent study, aiming to determine the role of various cognitive factors in the acquisition of two Russian cases, demonstrated further that L2 learners' metalinguistic awareness of rules is of vital importance in creating generalizations of inflection patterns, helping adult L2 learners encode information and form relevant declarative representations. Unlike in the previous study, the treatment included the production of case morphology, which might have contributed to improved generalizations, though this conclusion was not made by the researchers. Furthermore, the case inflection production was tested on new words, which were not included in the treatment (unlike

DeKeyser, 1997; and Roger (2011); see 3.2), thus, investigating both Function Assignment and Constituent Assembly. The researchers found that the accuracy of inflection production correlated with ability to generalize patterns from the input, which might possibly be interpreted as forming procedural representations. Also, the results of this study reinforce the benefits of deeper levels of processing (see 3.3.2) and could possibly be linked to 'noticing' (see 3.3.2).

Moreover, there are a few studies examining the effects of Processing Instruction (PI) (see 3.3.2.) specifically on Russian case acquisition, aiming to help learners consciously understand different functions of Russian cases and improve their correct encoding during input processing. For example, PI was tested for the effectiveness for the comprehension of Russian animate Acc., having delivered rather positive results (Morton, Yakimova, and VanPatten, 2011). Comer and deBenedette (2011) used PI for teaching learners to differentiate between directionality context, expressed by Acc., and locationality, expressed by Prep. Unlike the previous PI study, Comer and deBenedette's tests included a controlled production task, asking learners to complete sentences with an appropriate case form. The PI groups outscored the subjects, receiving Traditional Instruction on the interpretation tasks (testing encoding), but performed as well on the production task (testing retrieval), confirming Brooks & Kempe's (2013) conclusion that input processing activities need to be combined with oral production practice, which is also in line with the conclusions made in 3.3.3.

Another study, investigating the effectiveness of explicit instruction and, though without stating it directly, promoting 'noticing', is by Arnett & Lysinger (2013), who developed a set of materials containing pictograms of case functions (from Janda and Clancy, 2002) and tested their use for explaining case meanings within a standard Russian course. Their 17 participants were first year University students. The three test tasks included identifying a case in the input, filling the blanks and free writing, with oral production not tested, though it was included into the standard course. Arnett & Lysinger (2013) demonstrated significant difference in performance with regard to cases, between an experimental group and a control group, where pictograms were not used. This study illustrates the importance of explicit grammar explanations for beginner learners' performance on case.

Two very recent studies by Gor, Chrabaszcz, and Cook (2017; 2019) analysed how non-native speakers of Russian process case forms in the input. The researchers investigated the differences in processing Nominative forms (citation forms) and oblique case forms, as well as comparing the processing of overt and zero inflection. They found that, unlike native

speakers, non-natives do not necessarily display additional processing costs for oblique case forms or overt inflection, which means that L2 learners do not “decompose” inflected words in the input, unless required by the task. In addition, the second study demonstrated that proficient L2 speakers of Russian demonstrated sensitivity to ambiguous inflections. Though both studies examined the intricacies of word processing not relevant to the current study, their finding about learners not registering case inflection during recognition, supports VanPatten’s (2004) hypothesis of the primacy of lexical processing, arguing that L2 learners do not process inflection, unless specifically instructed (see 3.3.2). This reinforces the importance of ‘noticing’ for processing case suffixes in language input in an L2 classroom.

From the point of view of language pedagogy, as Cherepovskaia et al. (2021) noted, there are not many studies that investigate Russian case acquisition in classroom environment. The majority of studies investigating the effect of the type of instruction on Russian inflection acquisition (see 4.1.4) appear to opt for pedagogical experiments, conducted in laboratory conditions, when participants devote all their instruction time to learning particular forms (Comer & deBenedette, 2011; Kempe and Brooks, 2008; Brooks, Kempe & Sionov, 2006; Bowles, 2004; among others), sometimes even without being able to read or form sentences in L2 (e.g., Kempe and Brooks, 2008). Though the outcomes of these studies demonstrate what learners are capable of learning, I believe that laboratory conditions put certain limitations on classroom application of their findings. Ellis (2003) suggested that if we are “to make the shift from theory to practice it will be necessary to go beyond the psycholinguistic rationale…to address the contextual factors that ultimately determine what materials and procedures teachers choose” (Ellis, 2003: 337). That is why, one of the aims of the current study was to obtain results that would inform language teaching and would be generalizable to ab-initio courses.

I identified four studies that were conducted at different times and tested case production by L2 learners of Russian in instructional settings: Thomson (1980), Rubinstein (1995a), Arnett & Lysinger (2013) and Cherepovskaia et al. (2021). Though three of these studies (Thomson, Rubinstein and Cherepovskaia et al) did not investigate the effect of a type of instruction, all of them measured success rates for each of the Russian cases. The majority of these studies worked with English speaking learners of Russian, except Cherepovskaia et al (2021), who reported on Catalan-Spanish speaking bilinguals. Also, for the two earlier studies, students of higher proficiency levels were recruited from intensive programmes, while the two later studies tested beginners who completed standard University courses. It would probably be true to say that any direct comparison between the results of the current study and the above listed studies, would not be justifiable, due to differences in participant

pools, testing procedures and length of instruction, among other factors. However, the four studies discussed here, tested case production of L2 learners of Russian in instructional settings. Thus, it could be possible to look for some commonalities, which could be indicative of certain tendencies in case inflection accuracy. Table 4 presents a brief summary of these four studies.

Table 4

Brief Summary of Studies Investigating Russian Case Inflection in instructional settings

	Proficiency level	*Instruction period*	*Testing mode*	*Overall success rate*	*Success rate for Prep. and Acc.*
Thomson (1980)	1st year of an intensive programme	One year intensive course (hours were not specified)	Interview	Not reported	74% and 75%
Rubinstein (1995)	intermediate level after intensive study (two groups)	1410 – 1645 hours; 7 hours a day	Interview	60% for Group 1 (midcourse); 74% for Group 2 (endcourse)	69% and 67% for midcourse; 77% and 79% for endcourse
Arnett & Lysinger (2013)	Year 1 University course (hours are not specified)	15 minutes of explicit instruction when each case is introduced within a standard Russian course	Writing 5 sentences about a cartoon	63% (for the written production task)	Not reported
Cherepovskaia at el (2021)	A1 and A2 (as well as higher levels)	150 hours (two classes a week)	Writing a story within 30 min.	64% for A1; 75% for A2	92% and 53% for A1; 81% and 76% for A2 (calculated from the percentages of errors)

It is important to note that oral production of inflection (in an interview) was only tested with more advanced students (Thomson and Rubinstein), while Year 1 testees were given a writing task (Arnett & Lysinger and Cherepovskaia et al).

In addition, all, except Arnett & Lysinger, report success rates for each of the Russian cases. This split reveals an interesting observation with regard to the production of Prep. and Acc.. On one hand, Thomson (1980) and Rubinstein (1995), who tested more advanced learners, report similar percentages for the two cases - 74% : 75%; 69% : 67%; 77% : 79% respectively (with 1-2% difference). On the other hand, in the study by Cherepovskaia at el. (2021), who collected writing samples from Year 1 students, Prep. appears to have an advantage over Acc. in success rates, though the gap between the two decreases, with

growing proficiency levels, from 39% difference for A1 to 1% for C1 (see Table 5 for details). It appears useful to examine these differences further.

Table 5

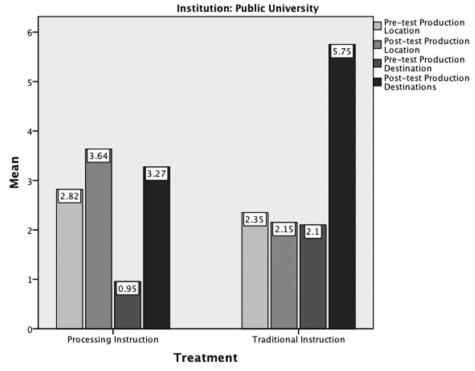
Prep. and Acc. Error Rates from Cherepovskaia et al (2021), p.16, with Calculated Success Rates.[2]

Level	Prep. Errors	Prep. Correct[2]	Acc. Errors	Acc. Correct[2]
A1	8%	92%	47%	53%
A2	19%	81%	24%	76%
B1	12%	88%	25%	75%
B1+	14%	86%	15%	85%
B2	13%	87%	12%	88%
C1	9%	91%	10%	90%

It is clear from the above data that, at the onset of learning, students have significantly more success in Prep. than in Acc. (92% and 53% respectively, for A1). At A2, the accuracy for Prep. decreases considerably (by 11% - from 92% to 81%), then going up to 88% at B1 and remaining at a similar level (± 1%) after that. Also, the participants' means bar chart from Comer & deBenedette (2011) (Figure 9) shows that at least some of the six Russian beginner groups in their study discussed above, displayed a similar pattern, when the success rates for Prep. came down, while the Acc. scores went up.

Figure 9

Bar Chart Demonstrating Means for Prep. and Acc. (from Comer & deBenedette, 2011)



[2] Cherepovskaia et al (2021) provide percentages of correct case inflections calculated from the totals of inflections for each *level*, rather than from the totals for each case (as in the other three studies). That is why, here, their correct scores are calculated from their error rates, which were computed by Cherepovskaia et al. from case totals.

This development appears worth investigating in the light of learners' processing, as the drop in the Prep. accuracy could possibly be caused by the introduction of the second case, that is, Acc.

Finally, none of the studies above investigate the effect of gender, case contexts or familiarity of lexis, which were identified in 4.1.2 as possible factors influencing the accuracy of case production.

To sum up this brief review, the experimental studies above support some of the findings, discussed in previous chapters, with regard to Russian case inflection. Moreover, it could be true to say that most of the studies are short-term experiments, mainly, except Gor et al (2017; 2019) investigating the effect of instruction on metalinguistic knowledge about Russian cases (that is, encoding), whether in relation to meaning or form. There are four studies that tested case inflection in classroom environment, namely, Thomson (1980), Rubinstein (1995a), Arnett & Lysinger (2013) and Cherepovskaia et al. (2021). They report a distinct tendency of changes in Prep. and Acc. rates at lower levels, when initially high accuracy for Prep. drops considerably at some point before levelling up with Acc. later. This appears to present an interesting opportunity to be investigated from the angle of learners' processing. Finally, none of the studies seem to explore different options of organizing the learning material and practice within the beginners' curriculum. There do not appear to be any studies investigating interactive practice or testing spontaneous production of case inflection in unfamiliar contexts either; the present study will fill this gap.

4.2. New L2 Grammar Teaching Framework, compatible with learners' processing and explained using Russian case inflection

In this section, I will propose and describe in detail my new grammar teaching framework, which is designed to address the inferences that have been made from the discussion of the relevant research in the previous two chapters (see 3.4), as well as those in the previous sections of this chapter, and which I argue could be able to assist L2 learners in acquiring grammar in speech more effectively, and potentially initiate proceduralization of some of the processes involved in acquisition (see 3.4.). In line with the main goal of the present study, the proposed framework aims to increase the effectiveness of grammar instruction by making it more compatible with the ways in which learners process information during language learning, thus improving the accuracy of the production of grammar forms and, consequently, reducing variability.

First, I will outline the main principles of the proposed framework and describe the system of presenting knowledge which lies at the basis of it. Then, I will apply these general principles to the teaching of Russian cases, which takes place in the first year of learning Russian, that is during Year 1/beginners' course, at the start of which learners are not expected to have any previous knowledge of Russian or any linguistic background. To start with, I will show the principal differences between the proposed framework and traditional grammar syllabus. After that, I will demonstrate how the new framework principles can be applied within one lesson at the very beginning of the introduction of cases. More specifically, I will be looking at optimizing the encoding of both explicit explanations and implicit information in the language input, as well as of the two types of retrieval of grammar forms, namely, recognition and recall, incorporating activities for learners to execute Function Assignment and Constituent Assembly, essential for accurate L2 speech production. In addition, I will be examining how 'noticing' and references to broader, more systemic grammatical concepts could be exercised throughout learning, which might have the potential to contribute to establishing mental representations of target features.

Moreover, the proposed framework is a system of continuous integration of different elements at every stage of a lesson, as well as in every part of a grammar curriculum, enabling learners to build their grammar knowledge simultaneously with developing their speaking skills, thus expanding their interlanguage. That is why, next, I will explain how different elements of information about the two initial Russian cases could be sequenced within the first 10 hours of teaching cases, for them to be "internalized" by learners, using cognitive resources that are available to them, illustrating it with concrete examples.

Finally, I will demonstrate how the same principles could be applied to other Russian cases, and will propose an alternative syllabus for a beginners' Russian language course, based on the the proposed framework (the learning outcomes for the proposed course can be found in 4.2.3). The main challenge here is to restructure the content and to organize classroom learning in such a way that could incorporate all the above processes in the most effective way and, at the same time, to attend to the constraints of a standard beginner language curriculum. The rest of this chapter explains how I addressed that challenge and gives reasoning for my choices and the decisions that I have made.

4.2.1. Spiral Curriculum (SC) and learners' processing

As a core for my framework, I have chosen the Spiral Curriculum (SC), which is a particular way of introducing learning material, put forward by the American psychologist and

educationalist Jerome Bruner (2009/1960), who laid the foundations of cognitive learning theory.

There are a number of reasons of why I identified SC as a basis for the new framework. First, SC provides a suitable alternative system for organizing learning within a curriculum, which appears to be more compatible with learners' processing than traditional curricula. Second, SC principles align really well with the seven inferences that I made in 3.4, allowing to segment the information for encoding and implement the idea of multiple retrievals, among other aspects. Third, Bruner (2009/1960) proposed his SC with leaners' cognitive development in mind and believed that learners actively construct their own knowledge, therefore, in my opinion, it has a good potential to facilitate the development of learners' interlanguage. Fourth, Bruner's work on SC is considered a landmark in conceptualizing learning and curriculum development and has become a crucial factor in the generation of a range of successful educational programs and pedagogical experiments. Today, Bruner's ideas have been effectively applied to design curricula in various subjects - sciences (e.g., Tytler, 2007), mathematics (e.g., Dean, 2010), medicine (e.g., Davis and Harden, 2003), engineering (e.g., Lohani et al., 2005), among others. Finally, the idea of applying the SC idea to teaching languages, first put forward by Howatt (1974) and Corder (1973) (who were instrumental to the field of applied linguistics), was developed further in a number of studies carried out in this field since. These mainly research teaching English as a Second Language (e.g., Crowley, 2022, Far, 2008; Kirkgoz et al., 2016), but there is an impressive example of the entire Maori language curriculum designed and effectively implemented in mainstream schools in New Zealand (Johnson & Houia, 2005). As far as my investigation could stretch, I could not identify any attempts to apply Bruner's SC theory to teaching inflection or to teaching Russian.

The main concept of spiralling involves information being structured in such a way that complex ideas can be taught at a simplified level first, and then re-visited at more complex levels later on; thus, learners are taught at levels of gradually increasing difficultly. Unlike teaching in a traditional "linear" way (Howatt, 1974), when each new point is studied as a whole and learners are expected to retain abundant details for future use, spiralling involves 1) introducing key concepts in a simplified form with minimum information first, and then 2) revisiting them later, 3) adding more details each time, to consolidate previous input (see Johnston, 2012, for a brief summary). This way, on each return to the topic, learners' knowledge and skills rise to a higher level (hence, the spiral analogy).

Later advances in cognitive psychology, particularly in memory studies and information processing, have created more sound theoretical underpinnings for SC. For example, the limited capacity of WM (Baddeley, 1992 – see 3.1.) points towards the validity of SC first principle of simplifying ideas and adding new information gradually. Also, it has been established that we retain information better if it is associated with the other information which is already stored in memory, (known as the phenomenon of elaborative processing which is remembering by linking to related information (e.g., Anderson, 2015)), confirming Bruner's proposal of connecting "new learning" to "old learning". SC second principle of revisiting feeds directly into the idea of incubation period when learners often perform better after having a gap in attending to a particular task (e.g., Anderson, 2015; Choi & Smith, 2005; Lightbown, 2008).

SC, initially proposed for early years learning, differed from the dominating (at that time) developmental psychology of Piaget (Bruner, 1997), which mainly followed linear structure, reserving complex concepts for later stages of learning. Bruner advocated introducing key concepts "in some intellectually honest form" (Bruner, 2009/1960: 33) at the start of learning, with the aim of building the entire structure of the subject on them and placing any subsequently introduced parts of information in a larger context, in order to enhance learner's understanding of the subject. This could be linked to the concept of deeper levels of processing, which is assumed to initiate conceptual thinking (see 3.3.2), and appears most relevant with regard to building learners' knowledge of L2 grammar, whether it is envisaged as "(re)-assembling" of a system of syntactic structures or "restructuring" a system of form-function connections, as discussed in 2.3.

The idea of non-linear teaching has been supported by Gass (2015), who criticises the opposition for their underlying assumption that L2 learning is linear and goes from one stage to another. Lightbown (2008) sees linear grammar teaching as "AAABBBCCC", illustrating it by the sequence of verb tense forms in the L2 English curriculum, each of which "disappears from classroom language" to be replaced by the next one (Lightbown, 2008: 40) (see Figure 11). As an alternative, she advocates "spacing", that is, distributing the learning of a particular element over time in a different pattern -"ABC, ABC, ABC", explaining its effectiveness by items being "more accessible" for retrieval next time round and by providing time "for the internalization, re-structuring and off-line processing" (Lightbown, 2008: 41). I believe that Bruner suggested the approach that goes further – in addition to retrieving the information that has already been studied, it offers to put that information in a new context and to add new information to it. This way, the pattern would look like something like this -

“A, AB, BC, ABC”, which would allow all the elements to get through the “bottleneck” of WM, in order to be firmly encoded and successfully retained in LTM. Without knowing all the intricacies of information processing and functions of different types of memory, that have been made available to researchers since 1960-s, Bruner proposed a system of organising learning in a way that appears to be more compatible with the way in which our processing works.

Furthermore, for Bruner, an important part of the required system, along with “knowing that”, is “knowing how” (Bruner, 1965), which echo the distinction between declarative and procedural knowledge defined in 3.2. As Bruner phrased it, “learning subject structure involves supporting habits and skills that make it possible for pupils to make active use of the materials that they have come to understand” (Bruner, 2009/1960: 12) . Thus, the building of the system of skills is inseparable part of learning and is connected to building the system of knowledge about facts; this being directly in line with the idea of proceduralization. At the same time, the idea of growing knowledge and growing system loops back to de Bot’s definition of language acquisition which was adopted in 3.3.3 (de Bot, 1996: 531). Furthermore, in connection with the two types of knowledge, Bruner differentiates two types of “transfer” of knowledge – “specific”, which can be summed up as transfer of the learnt skill to similar tasks; and “non-specific”, which is explained as an application of acquired knowledge in different situations (Bruner, 2009/1960: 17), which goes beyond proceduralization and is closer to a decision making process required for selecting correct grammar features (e.g., case function) during Function Assignment (Figure 2 in 3.1), involving the assessment of the context.

It appears that Bruner’s SC offers an effective system of structuring learning in educational contexts, which has been well-attested for various subjects, including some languages, and which creates conditions for optimizing learners’ processing, which is the main focus of the present study.

4.2.2. Encoding: processable amounts and simplified concepts

If we are to follow SC principles and to make inflection processing more effective, the amount of explicit grammar information, traditionally presented to language learners at one time, for example, within one lesson (see example in Figure 4 in 4.1.2), is to be divided into smaller “portions” which would be introduced in a few steps, with each of these steps representing a “coil” (or a “turn”) of the learning spiral.

Also, the first case, which is presented to learners needs to be as easy to process as possible, as well as suitable for establishing a firm initial concept of case. Textbooks normally opt for Prep. (e.g., Kudyma et al, 2015) or Acc. (Aizelwood, 2007). I argue that, initially, Acc. is considerably more complex for learners than Prep. for the following reasons: a) Acc. differentiates between genders (has separate suffixes), as well as marking animacy and having two allophones in feminine (see Figure 3 and Table 3). b) Acc. has no overt inflection (zero suffix) for masculine Singular, which is sometimes seen as an advantage, but it is syncretic with Nominative and would be very unhelpful in establishing the concept of case and, consequently, creating representations (thus affecting Function Assignment during production), due to the lack of transparency between the two case forms. c) Acc. is often used with no prepositions, e.g., in patient function (prepositions serve as a strong cue in case assignment (see Comer & deBenedett, 2011)). These three aspects of Acc. demand additional cognitive resources. Among the books evaluated in 4.1.3 (See Appendix A), Langran and Veshnyeva (2012) choose to introduce Genitive Singular first, which appears to have even more processing challenges, than Acc. Though suffixes for Genitive Singular are overt, masculine nouns are inflected with “a” (or â) which is syncretic with the feminine base form (See Tables 2 and 3) and would be very confusing for beginner learners. Also, to inflect feminine nouns for Genitive, the 7-letter spelling rule has to be learnt. Furthermore, Genitive case functions are notoriously diverse and often are quite abstract (e.g., negation or possession), which would make Function Assignment quite problematic. In addition, there is an issue of countable/uncountable nouns, which require different inflection in Genitive and are not always the same in English and in Russian (for example, “fruit” in Russian are countable and can be Plural, while “potatoes” are not and are always Singular).

In contrast, the main suffix for Prep. (“e”) is the same for all three genders and is different from their base forms, making both Function Assignment and Constituent Assembly fairly straightforward. Importantly, Prep. always follows a preposition (*v* normally for “in”, *na* normally for “on” and *o* for “about”), which, at the initial stages of learning (before other meanings for the same prepositions are introduced) is a distinct cue, assisting in selecting a case function at this stage. That is why within my proposed framework, Prep. is taught first.

Within a traditional linear Russian language syllabus, five Prep. Singular nominal suffixes for three noun classes and three subclasses - “e”, “ii”, “i”, “u”, “û” (see Table 3) are normally presented in the first lesson and practiced for a number of consecutive lessons. (Plural case forms are normally introduced separately at a later stage and, therefore, are discussed later.) Often, adjectival declension, which has an additional set of four suffixes for Prep.

Singular, differentiating between genders - “om” and “em” for masculine, “oj” and “ej” for feminine, is introduced at about the same time (e.g., Robin, 2011). Normally, learners would be expected to memorize the entire nominal paradigm (often together with the adjectival paradigm) by spending hours practicing the full set of suffixes (which represents passive retrieval – see 3.3.1) and, then, weeks, or even months, trying to “internalize” (or proceduralize) that knowledge.

Within the proposed framework, following the first SC principle of introducing information in smaller amounts, learners would deal with one suffix at a time, which would allow them to process it from a short explicit explanation within the scope of available cognitive resources and transfer it for storing into their LTM. This would be then followed by encoding it during recognition from language input, e.g., reading (with ‘noticing’ encouraged), and then recalling it in short meaning-focused interactions, getting ready for the next grammar form. (Adjective declension is introduced at a later stage and is discussed further.)

In practice, this would mean that teaching the Prep. case would need to start with explicitly explaining only one function (e.g., a place/location, which is apart from being most frequent, is pragmatically obvious) and one noun suffix (e.g., “e” for most singular nouns, Class 1 and Class II). This would constitute one “portion” of learning, producing forms like *v London**e*** (meaning “in London*[Prep. masc.]*”). Then, in SC terms, the first “portion” of Prep. has to be put into the context of “old learning”, in here, that of the Nominative case with its Subject function (for learners to encode and retrieve the two functions and map them onto the two forms). Next, the new suffix would need to be ‘noticed’ at the end of masculine place names (not requiring translation) and, after that, in the inflected forms of familiar vocabulary. This would be contrasted with base-form zero-suffix (for masculine), to attract learners’ attention to the connection between function and form (recognition of form in language input). All case forms need to be read aloud to initiate passive Constituent Assembly (ex. 3a and 3b in Appendix E) before learners asked to assemble them independently (ex. 4a). A small sub-rule for monosyllabic Class 1 nouns (see 4.1) is introduced to add a different inflection (“u”), which is learned as a chunk to start with in *v sadu* (for “in the garden”). Then, case forms are put in context of short phrases offered for reading (ex.4b in Appendix E). (Note that this material is normally used around the 5th - 10th hour of complete beginners’ course, depending on whether learners are expected to master Cyrillic independently before the start of the course, which is common in some UK Universities). The focus now is on meaning and learners are tasked to remember where things and people are, to enable them to answer questions in controlled production exercise 4c, when both form and meaning will be active. ‘Noticing’ would be encouraged throughout. Finally, this “coil” of the learning spiral finishes with an

interactive task-based activity in 4d, when learners work in pairs and have to establish where their partner's people are, using their randomized list of people and places. This would involve unprepared short contextualised interactive exchanges, with some results reported back to the class, where all stages of inflection encoding and retrieval would be performed, including Function Assignment and Constituent Assembly. Figure 10 visually illustrates the cycle of multiple instances of encoding and retrieval (see 3.3.3.).

Next, the same Prep. suffix ("e") is practiced with feminine nouns. As discussed in 4.1.2, feminine case forms are assumed to have one more step in their Constituent Assembly, when the feminine base-form marker (e.g., "a"/"â") is removed, before the case marker (here "e") is added. This is likely to be more cognitively demanding and, therefore, needs to be practiced separately, to increase the effectiveness of acquisition, at the same time contributing to establishing gender differentiation.

Figure 10

"Recycling" Case Information Within a Spiralling Lesson (or a Section of a Lesson)

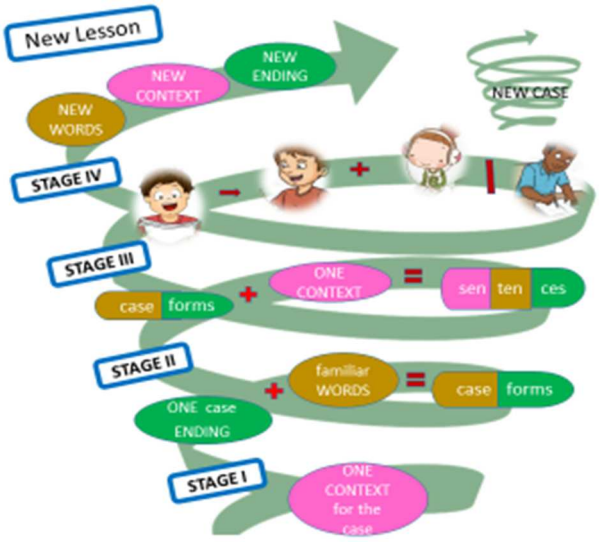


Figure 11

Grammar Topics Within a Linear Curriculum

If a course syllabus has a limited number of hours, further recognition during listening and recall during writing might be given as homework. That is where 'noticing' becomes invaluable, shifting learners' focus fairly regularly from meaning to form and back. This kind of "recycling" of the inflection (M. A. Martin, 1978: 152) will enable learners to establish a skill of switching between Nominative and Prep. suffixes, thus selecting a correct case function

during the Function Assignment stage (see Figure 2 in 3.1.). New place names or other vocabulary can be fed into the lesson, in order to provide Constituent Assembly practice for less familiar vocabulary (Bruner's "specific transfer"). As discussed in 3.3.3, this should strengthen the mapping between form and function, which could potentially work towards creating a procedural representation of the form. Therefore, I am going to argue that following the sequence of short explicit explanation, small amounts of input processing and short interactive speaking practice in class, would provide the opportunity for running the full cycle of encoding and multiple retrieval necessary for developing increasingly automatized case form production, also addressing Function Assignment and Constituent Assembly.

According to Bruner (2009/1960), this ability to perform a particular procedure, then, can be transferred (Bruner's "specific transfer") to a similar task – here such a task would involve a different case suffix. The short processing cycle, described above, would need to be repeated for every case suffix and for every case function, with more similar practice suggested for homework. This ensures systematic integration of the teaching of grammar and the teaching of speaking, creating conditions for input and output processing. The discussion in Chapter 2 and Chapter 3 suggests that organizing learning in this way has the potential to optimize learners' processing and help learners acquire well-established oral production skills, combining linguistic and communicative elements.

However, spiralling does not just reduce the amount of information presented to learners at a given time – it is essential that different parts of grammar information are interconnected, so that learners' interlanguage is systematically "upgraded". Thus, each new portion of information is introduced not in isolation but as part of the whole, and is linked to what has been previously learnt, reemphasising the core concept. That is why introducing key concepts (in a simplified form) at the start of learning is paramount for SC teaching and is now linked to deeper learning, the effectiveness of which is, to a degree, empirically confirmed by psycholinguistics (see the discussion in 3.3.3 and 3.3.4).

Developing my Prep. case example further, the first case suffix that learners practice ("e"), needs to be put in the context of the Russian inflection system (or "system of suffixes" for learners), required to keep words connected within the flexible word order, as, in Russian, the words do not have a fixed position in a sentence, which would suggest a function, as, for example, in English. This would create a basic concept of inflection that relates to the whole L2 system and consequently, can then be transferred to other cases, gender agreements, verb conjugation and other parts of the system (Bruner's "non-specific transfer"). This is why,

earlier on, I argued that Prep. needs to be introduced first as its fairly transparent forms would assist learner in understanding the functionality of inflection from the start.

This way declarative knowledge about the language system will continuously expand, while regular interactive practice will expand the system of procedural skills. The existence of two parallel systems - one of mental grammar and another one of production skills has been speculated by Wong and Van Patten (2003); and is also discussed in 3.2. Within the proposed framework, the core grammar syllabus and the communicative syllabus become intertwined, continuously complementing each other without overriding priority given to either. This synthesizes previous developments in integrating the teaching of grammar with the teaching of speaking, discussed in Chapter 2, and makes the proposed framework pedagogically innovative, presenting an optimal way of organizing a language curriculum. The exact workings of the Russian Beginners Case syllabus built using the proposed framework, is discussed in the next section.

4.2.3. Multiple retrieval: revisiting – not revising

Another crucial difference between spiralling and linear teaching is the principle of revisiting, which is not at all the same as revising. It is true that in both cases, learners come back to what they have learnt, but unlike revising, which goes over the same material, revisiting builds on previous knowledge and expands it by adding more information, connected to what was learnt before.

For example, when the first Prep. Singular suffix “e” (introduction of which was discussed in the previous section) is revisited, a different Prep. context for the same suffix is explained (e.g., time references requiring Prep., e.g.: *v mart**e*** for “in March”), or a different Prep. suffix is introduced (e.g., “ii” for “iâ” subclass, e.g., *v Ross**ii*** for “in Russia”), or a different preposition requiring Prep. case is presented (e.g., *o* for “about”). Thus revisiting builds up on the initial material by connecting more information to it; as opposed to introducing all five suffixes at the start, then practicing the full set throughout the topic and revising the same five at the end, as in linear curriculum. Following this logic, I am suggesting that a learning spiral could be envisaged as an inverted conical helix (spiral) rather than a cylindrical coil, representing an expansion of knowledge and skills. This is rather dissimilar to linear learning when topics can be represented by the same size blocks stacked on top of each other, with amounts of information being similar at the start and at the end of a block (see Figure 10 and Figure 11). Moreover, revising is normally planned for the end of a topic, while revisiting is intermittent with other grammar topics and activities.

Furthermore, revisiting allows to introduce grammar in such a way that would assist in gradual building up of beginners' speaking practice, expanding learners' active database of structures and vocabulary, enabling them to interact by creating more and more sentences of their own. At the same time, these interactions provide conditions for retrieving more varied inflection patterns (rather than producing prefabricated chunks), as well as for exercising 'noticing' and modifying their output, thus activating the connections to various representations. For example, the original Prep. Singular noun suffix "e" can be followed by 1st person Singular present verb suffix "û", thus creating an ample base for beginners to talk about where they work, live or play (e.g., "*â rabotaû[1st person Sing.Present verb] v bibliotekę[Prep.fem.noun]*" for "I work in the library."). Then, the Prep. noun suffix becomes "old learning" for the new explicit "portion" of verb forms and provides the context for more interaction, which later could be developed by learning other verb forms.

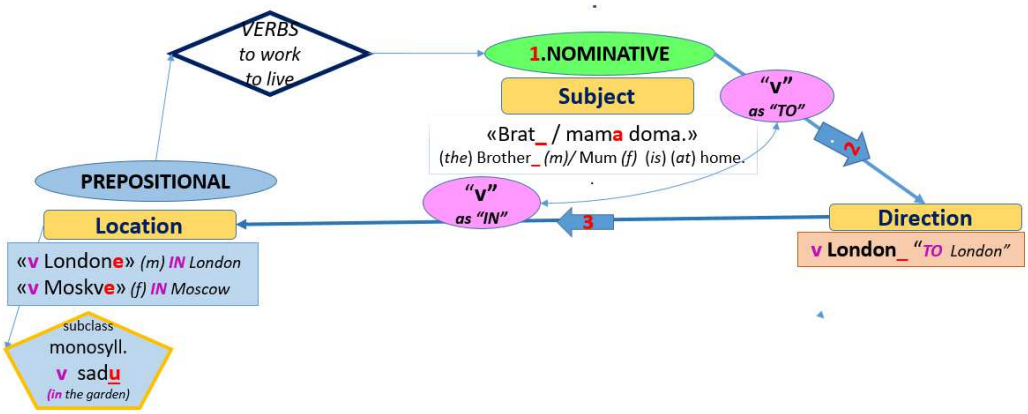
The next step would be Acc. case, which would create an opposition of two oblique cases and, at the same time, introduce a new topic of travelling, involving directionality context (see Comer & deBenedette, 2011, discussed in 4.1.3), e.g., "*Â edu[1st person Sing.Present verb] v London[Acc.masc.noun].*" for "I am going to London". This would create new challenging conditions for using Prep. case forms, as the same places can be used after the same prepositions (see 4.1.2 for details) and a case is determined by pragmatic context, often prompted by a verb. The "ii" suffix (for "iâ" feminine subclass) can be introduced later when learners are studying Past tense, as it would bring in the names of various countries (ending in "iâ"), thus revisiting the Prep. case at a different level, when the challenge is to assign a different suffix for the familiar case context. This way the whole beginners' curriculum can be spiralled, incorporating smaller spirals of cases and verb conjugations, as well as even smaller "recycling" spirals of individual lessons (see further discussion of learning materials in 5.3.2).

A series of figures below (see Figures 12a – 12f) illustrate the exact sequence in which case forms and case functions are introduced into learners' production during ten hours of learning from the moment the first case suffix is presented. Following the spiralling principle of revisiting, case contexts for Prep. are alternated with those for Acc. The order of introduction of case contexts is determined by several factors, with ease of processing and compliance with spiralling principles as the absolute priority, but usability in structures handled by learners and relevance to learners' pragmatic contexts, as the next most important criteria.

As it can be seen from the Figure 12a, at the start of case introduction, the opposition of the presence of inflection (in Prep. on the left) and absence of it (in Acc. on the right) after the same preposition *v* (for “in” or “to”) is presented, in order to attract learners’ attention to the importance of case suffixes and emphasize their connection to meaning. However, initially, only Prep. is practiced, to ease the processing demand. In addition, a chunked monosyllabic phrase *v sad**u*** (for “in the garden”) represents a Prep. subclass, which has a different suffix, but is not practiced on other vocabulary.

Figure 12a

The Order of Introduction of Case Suffixes and Case Contexts (see stages 1-3)



Next (see Figure 12b), a different preposition for the same (Prep.) case is introduced – *na* (used in the meaning of “on” here), providing more practice for the same Prep. suffix “e” and for the same location context. Learners’ task here is to select a correct preposition – *v* or *na*.

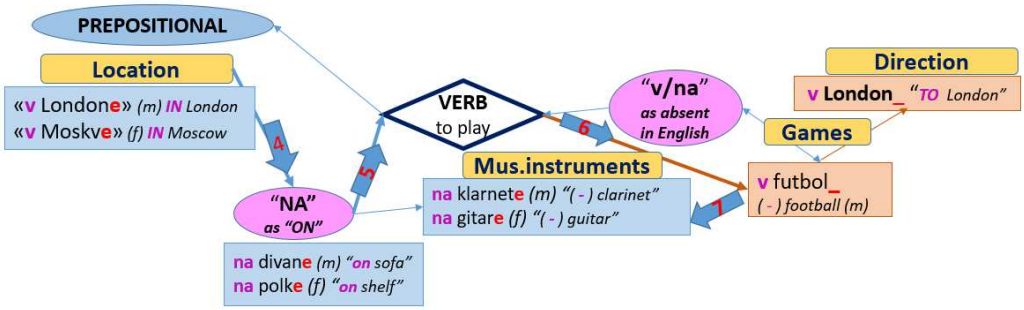
After that, sports games are learnt - these are used after the same preposition *v* which normally stands for “in” and is absent in English in front of games. Importantly, this is an Acc. context, and, as the absolute majority of sports games in Russian are masculine, they have the zero suffix (that is, do not require an inflection). The purpose of the introduction of games at this stage is twofold – first, this trains learners to exercise ‘noticing’ in the input, as well as in the output, (that is, notice the presence or absence of the suffix, e.g., *Â igraû v futbol_ v park**e***. for “I play football in the park.”); second, they learn to differentiate between different case contexts for the same preposition, but the distinction is clearly indicated by the verb “to play”, which is transparent from the processing point of view.

The next step (still in Figure 12b) is to teach musical instruments, as they are used after the same verb (*igrat'* for “to play”) but require Prep. (as opposed to games, which needed Acc.),

as well as a different preposition *na* (which normally stands for “on” but is absent in English in front of musical instruments). This introduces the idea that the same verb can be followed by prepositional phrases in which different cases are used. Nouns for musical instruments can be masculine or feminine, but in Prep. both genders take the same suffix (“e”) in the majority of nouns, so learners just learn to differentiate between “e” for Prep. and “no “e”” (that is, zero inflection) for masculine games in Acc.

Figure 12b

The Order of Introduction of Case Suffixes and Case Contexts (See Arrows 4-7).



In Figure 12c, the first new information is the use of the familiar prepositions (*v* and *na*) in the new meaning of “at” (in Russian, there is no preposition directly corresponding to “at”). This does not affect case inflection at this stage, but is essential for learners’ interaction, as well as for up-coming directionality context. However, the main new challenge here, is the feminine Acc. suffix (“u”), which, initially, is used only with days of the week. These belong to the closed lexical set – there are 3 masculine, 3 feminine and 1 neuter (which remains unchanged in Acc.) words in this set, which are frequently used in conversations, often becoming chunks.

The next big step is the use of already familiar Acc. suffixes (used for days of the week), both masculine and feminine, in the new directionality context. This requires differentiation between the location context (Prep.) and the direction context (Acc.). This is considered one of the biggest difficulties for learners with regard to cases (e.g., Comer & deBenedette, 2011) and is very challenging for cognitive processing, as both forms can be used with the same vocabulary (places and place names) and after the same prepositions (*v* and *na*). The only cues are the pragmatic meaning of being stationary (for Prep.) or moving (for Acc.), normally aided by verbs (for example, verbs for going or travelling would need Acc.). Producing sentences with directionality context also involves differentiating between genders (which was not required in the Prep. location context), as well as employing a fairly new

overt Acc. feminine suffix “u”. This new context creates a new speaking situation of travelling, which is popular with learners, and allows later on to introduce vocabulary for transport and months, both revisiting Prepositional but using different prepositions (*na* and *v* respectively) - see Figures 12c and 12d below.

Figure 12c

The Order of Introduction of Case Suffixes and Case Contexts (See Arrows 8-12).

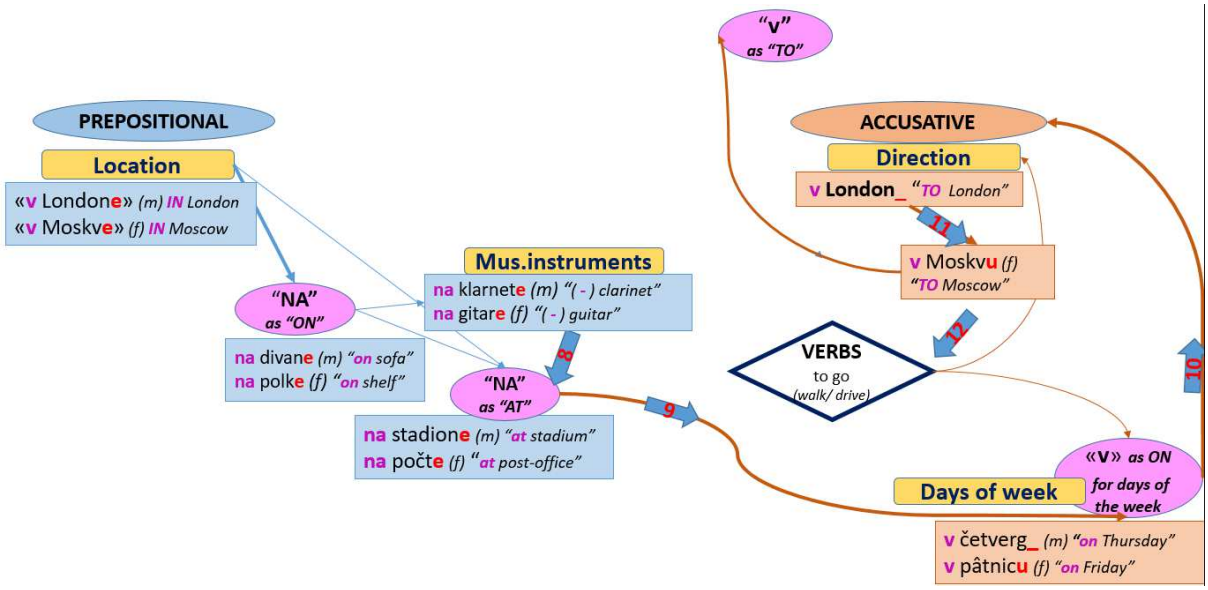
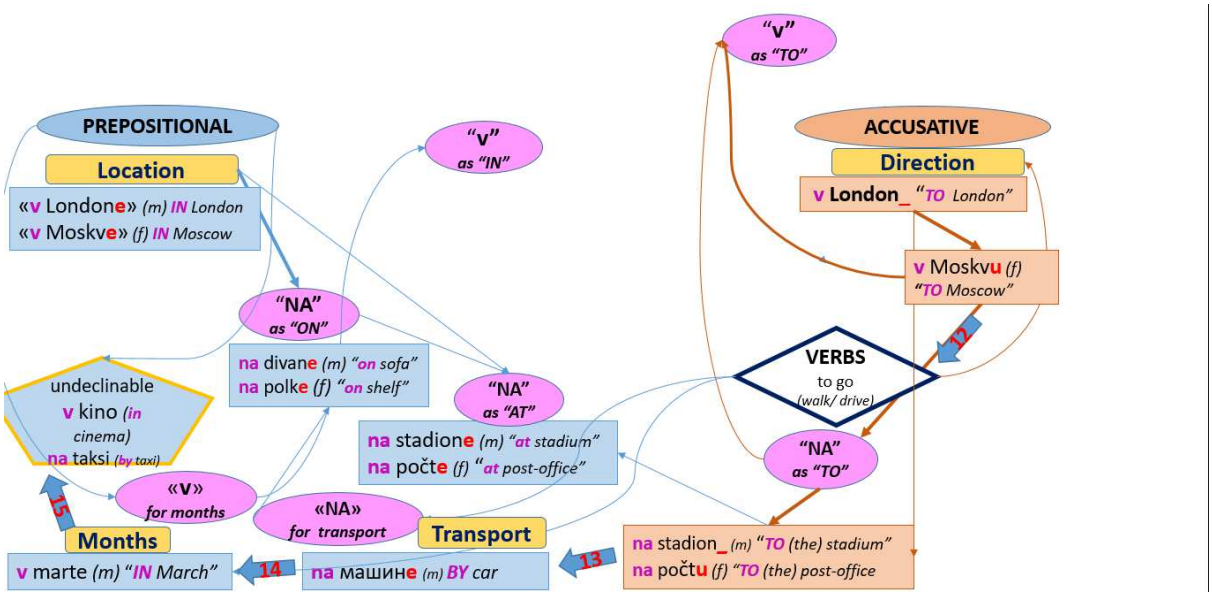


Figure 12d

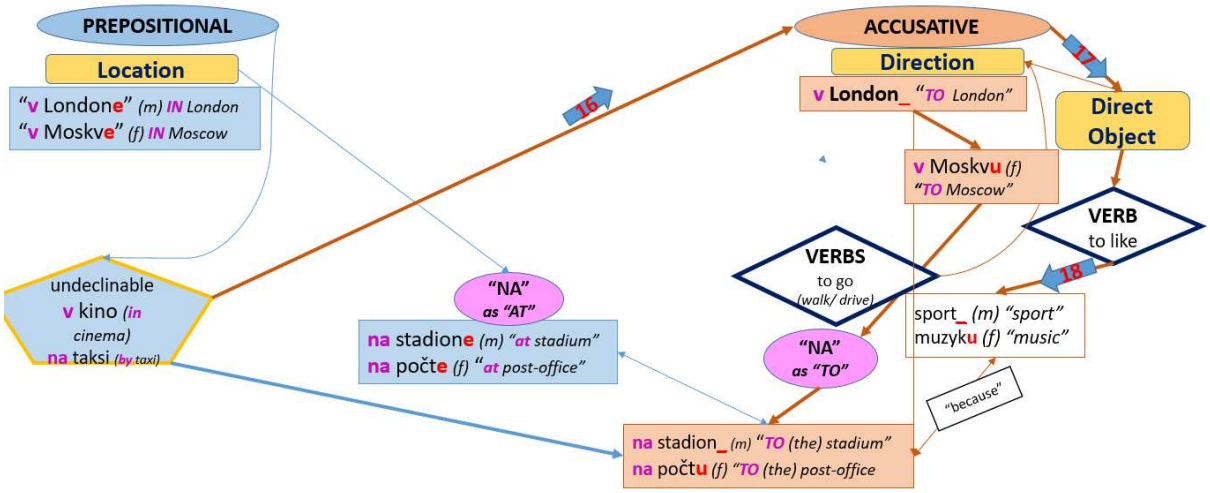
The Order of Introduction of Case Suffixes and Case Contexts (See Arrows 12-15).



At the end of the initial 10 hours of case learning, the indeclinable nouns are mentioned, which stay the same in both Prep. and Acc., and the idea of direct object function for Acc. (not requiring any prepositions) after the verb “to like” is explained. That allows learners to revisit the games and musical instruments introduced earlier, as well as linking to the topic of going (to concerts, exhibitions or stadiums), using the Russian conjunction for “because”, thus developing syntactic structure by embarking on subordinate clauses (see Figure 12e).

Figure 12e

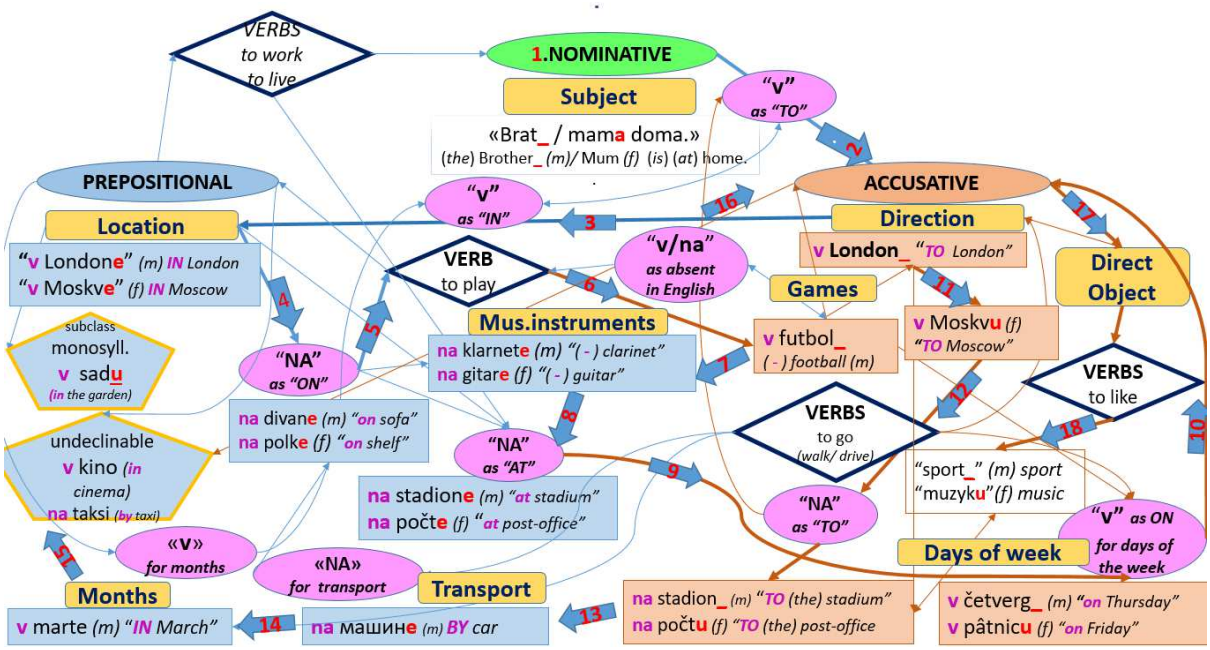
The Order of Introduction of Case Suffixes and Case Contexts (See Arrows 16-18).



The whole scheme of introduction of different segments of information about Prep. and Acc. (case suffixes, case contexts and prepositions), gradually feeding these in, while shifting learners’ focus from one case to another, is presented in Figure 12f.

Figure 12f

The Overall Order of Introduction of Case Information Showing Range of Connections Built by Learners for Prep. and Acc.



4.2.4. *New proposed spiralling syllabus for teaching Russian cases within a complete beginners' course*

Following the same principles of my proposed framework in the way demonstrated above, and using my considerable teaching experience, I designed a full syllabus of teaching Russian grammar for a complete beginners' Russian language course, where nominal and adjectival case infection (Singular and Plural) is gradually introduced into learners' production throughout the first year of learning. As a basis for this syllabus, I used the module descriptions for two Russian courses at the University of Leeds - Beginning Russian 1 (SLAV1010), which runs in Semester 1, and Beginning Russian 2 (SLAV1010), which runs in Semester 2 (see here https://webprod3.leeds.ac.uk/catalogue/dynmodules.asp?Y=202223&M=SLAV-1010; https://webprod3.leeds.ac.uk/catalogue/dynmodules.asp?Y=202223&M=SLAV-1020). Similarly to the above courses taken together, the proposed spiralling course is planned for five contact hours per week during two semesters, each with 10 weeks of teaching, thus making 100 one-hour lessons in an academic year.

The objectives of the Semester 2 course (Beginning Russian 2) state that "the completion of the course" would allow students "to speak and write effectively on a number of structured topics studied during the year". Formally, the objective of the proposed spiralling course is similar with regard to acquiring the above language skills, as its *main goal* is to enable

beginner learners to utilize Russian grammar effectively and more accurately in their production. However, the main focus of the proposed course, reflecting the overall aim of the present study, is on *optimizing* the teaching of Russian cases, increasing the accuracy of their production and reducing the variability among learners. Therefore the paramount difference between the traditional and the proposed syllabi is in *restructuring the content* of a standard beginners' course in order to enable learners to process grammar information in the most optimal way and to accurately implement it in their production.

More detailed *learning outcomes* for the proposed course are as follows: By the end of this Year 1 course, students should be able to construct their own sentences in their spontaneous production, using all six nominal cases and verbs in three tenses within a simple clause structure, as well as using subordinate clauses, without relying on the pre-learned phrases and sentences. In addition, the majority of learners should be able to use adjectives in a full range of case forms (for the three genders) in certain contexts within the topics studied, as well as some Perfective Aspect verb forms. Moreover, students should be able to pass all standard written grammar tests and a relevant oral exam, approved by the HE for Russian modules for Year 1 for language degrees, with high grammatical accuracy.

The proposed restructuring of the standard contents has been achieved by systematically 1) segmenting grammar information for each case into processable amounts, 2) spacing it out throughout the course and 3) revisiting it at intervals, adding new information every time. The full syllabus is presented in Appendix L.

First, the segmenting of the information about Russian cases resulted in identifying an "initial" suffix, or a set of "initial" suffixes for each case (see Table 9 below). This is similar to how "e" was selected to be introduced first for Prep. Normally, a set consists of two "initial" suffixes – one for masculine and one for feminine, as neuter, as a rule, has suffixes identical to masculine (except Genitive Plural, where both feminine and neuter lose their vowel base-form marker). A different number of "initial" suffixes are allocated for only two cases - Prep., where both genders are inflected with the same suffix "e", and Genitive Singular, where feminine case marker is subject to the 7-letter spelling rule, determining the choice of the allophone ("y"/"i"), crucial for the inflection process; this resulting in a set of three suffixes introduced one after another within one lesson. Usually, no subclasses or allomorphs are introduced initially (except Genitive Singular, as explained above). Table 6 presents the "initial" suffixes for each of the six cases (plus Genitive Plural, which has a set of its own, due to the complexity of form) in the order they appear in the full syllabus in Appensix L.

Table 6

"Initial" suffixes for each of the six Russian cases

	Feminine	Masculine
Nominative	-a / â	ø
Prepositional	-e	-e
Accusative	-u	ø (inanimate)
Instrumental	-oj	-om
Genitive Singular	-i / y	-a
Dative	-e	-u
Genitive Plural	ø	-ov

In Appendix L, the introduction of "initial" suffixes is indicated by the case name highlighted in blue – this marks the first time when a particular case is explained to learners and appears in their production. There is no fixed interval and the choices for the turn of "initial" case forms are dictated by several criteria – the complexity of the inflection form for learners' processing, the complexity of case contexts, suitability for learners' language base at the moment of introduction, what other new material is introduced (e.g., vocabulary or verb forms) and sometimes even the time in the semester when the form is introduced. This is discussed further on in this section.

Similarly to "initial" suffixes, from the long list of case functions and case contexts, "initial" case contexts have been selected for each case and are presented in Table 7. These are matched with initial suffixes in Table 6 and are also listed in the order they appear in the syllabus.

Table 7

"Initial" case contexts for each case

Case	*"initial" context*
Nominative	Subject
Prepositional	a place (location)
Accusative	days of the week
Instrumental	"with"-constructions (an adjunct), e.g., "with a friend"

Genitive Singular	“of”-construction, e.g., “a glass of juice”
Dative	to go to see/visit people (*idti k* + animate)
Genitive Plural	use after *mnogo* (for “a lot of”)

The main criteria for selecting an “initial” case context is its transparency, in order to make it easily identifiable for learners, as it was explained for Prep. and Acc. in 4.2.3 (location and days of the week, respectively). As it can be seen from Table 7, these two cases are followed by Instrumental. In a traditional beginners’ course, Instrumental is often studied in Semester 2 and is considered a complicated case as it is usually used with reflexive verbs, which are complex to acquire (for example, Kudyma et al (2016) introduce instrumental in Chapter 17 from the total of 24 chapters). However, if we take the “initial” case context selected for Instrumental in Table 7, namely, phrases with the preposition “with” (e.g., *Â živu s bratom[Instr.masc.].* for “I live with my brother”), it does not require reflexives and is easily identifiable by the associated preposition *s* (for “with”) . At the same time, the “initial” case suffixes for Instrumental are straightforward and easy to process, as they are not syncretic with any suffixes accumulated by learners by that stage. Therefore, in the proposed syllabus, Instrumental is introduced in Session 24 from the total of 100 lessons (see Appendix L). By then, the main forms and contexts for Prep. and Acc. are established in learners’ production and Instrumental provides a large extension to learners’ speaking situations, going really well with recently introduced Plural verbal forms. Instrumental use with reflexives is studied later (Session 38).

Then, Genitive Singular appears towards in the second half of Semester 1 (session 32), nine sessions after the “initial” Instrumental – this is due to the complexity of its forms and functions (see discussion in 4.2.2). The three main challenges for learners in the production of Genitive Singular are – 1) Genitive masculine case marker is syncretic with feminine base-form marker (“a”), which normally causes a lot of confusion for learners; 2) 7-letter spelling rule for feminine allophonic suffixes, which cannot be easily scaffolded; 3) the array of Genitive case contexts is often abstract and difficult to identify. Genetive case context (“of”-constructions) is selected as “initial”, despite having no preposition (for “off”) in Russian, due to incredible frequency of use.

Next, Dative is reserved for Semester 2 (Session 59), as it has two “initial” suffixes, both syncretic with previously studied inflection (Dative masculine “u” / “û”, which is the same as Acc. feminine; and Dative feminine “e”, which is the same as “initial” Prep.). So the main challenge here is Function Assignment, which is also very complex in Dative, as in English Russian Dative is often expressed by Acc. or Nominative. The “initial” case context “*idti k* + animate noun” (see Table 10a), despite having a preposition which would normally ease processing in other contexts, translates into English with a different syntactic structure, involving a second verb - “to go *to see/visit* a person”, which is absent in Russian. Moreover, it is confused with Acc. directionality context (e.g., “to go to London”), as both have a meaning of “moving to” and follow the same verbs, as well as the same preposition in English. Other Dative functions are equally complex, but have a lower frequency than the one selected.

Finally, Genitive Plural is treated as a separate set of case suffixes and was assigned its own “initial” set of suffixes (for session 71), as Genitive is the only case in Russian that differentiates between genders in Plural, that has a neuter suffix identical to feminine (rather than to masculine, as in all other case sets) and, finally, many feminine and neuter case forms, when dropping the vowel base-form marker, have to have a vowel insertion between the last two consonants (to follow the rules of phonotactics), while the choice of inserted vowel is subject to the 5-letter spelling rule. In addition, Genitive Plural has three clusters of nouns which are inflected with a “suffix” (“ej”), which is different from the “initial” case set and is not an allomorph. All the above factors, coupled with the complexity of case contexts (explained in the Genitive Singular paragraph above) and the necessity to differentiate it from Genitive Singular, especially on the background of un/countable distinction, make Genitive Plural one of the most challenging categories to acquire.

As for other suffixes for the same case (allophones, subclasses and Plural), these are gradually introduced throughout the course, thus increasing the amount of inflection information that learners handle. Often, the complexity of forms and case contexts increases when the same case is revisited, but sometimes forms with lower frequency of use, are introduced later, when learners have a larger language base, allowing them to use these. Spacing between different forms of the same case varies considerably and is dictated not only by the gradual increase of cognitive demand but also by the frequency of use, as well as by the presence of other new grammar forms, e.g., verbal inflection, and appropriateness for speaking topics. For example, Plural for Prep. (“ah”) is introduced 15 sessions after the “initial” suffix (session 21), as, in that period of study, learners handle Acc. (both masculine and feminine), as well as conjugation of verb in Present (six suffixes altogether) and

adjective gender agreement in Nominative (for the three genders). Also, the frequency of Prep. Plural is not very high. In contrast, for Instrumental, Plural suffix is given in the session following the “initial” suffix (see session 24 and 25), as it is required for learners to speak about whom they go on holiday with and “with friends” is a popular choice. From the point of view of form, all three Instrumental suffixes (masculine/neuter, feminine and Plural) are transparent and do not present any reasonable processing challenges. Moreover, Instrumental is the third oblique case in learners’ production, thus the concept is already familiar. However, Genitive Plural appears in the syllabus considerably later (session 71) than Genitive Singular (Session 32). This is due to the extreme diversity of Genitive functions and contexts, in which it is used, which take time to gradually feed into learners’ production, and Genitive Singular needs to be very well-established in learners’ interlanguage before Genitive Plural is tackled, as learners would have to differentiate between the two sets of complex suffixes for the same functions. This is also complicated by countable/uncountable distinction, which is not the same in Russian and in English.

At the same time, learners’ growing ability to process similar grammar information faster (see ACT-R model (Anderson & Lebiere, 1998) discussed in 3.2) is taken into account. For example, Acc. feminine inflection is initially only represented by “u” (Session 11) and its allomorph “û” is presented in Session 23. However, for Dative which is introduced in Semester 2 and has a syncretic suffix (“u”) for masculine, the gap between the “initial” suffix and its allomorph is not as big, as the concept is already familiar to learners and they can transfer this knowledge to the new case.

As for adjectival case suffixes, initially, these are learnt in chunks, for example, *na prošloj nedele* (Prep. feminine, meaning “last week”) in Session 22 or *každuû nedelû* (Acc. feminine, meaning “every week”) in Session 29. These suffixes are practiced with other vocabulary later, as at the start of learning, learners sometimes struggle to differentiate between nouns and adjectives, which do not have an overt marker in English. However, adjectival Acc. feminine suffix “uû” is used with other adjectives first (Session 37), as it is unique (not syncretic to any other suffixes) and can be directly associated with the Acc. feminine nominal suffixes “u”/ “û”, thus easier to process. On the contrary, “initial” adjectival suffixes for Prep. (“om” and “oj”) are completely different from nominal Prep. suffixes (e.g., “e”, “ii”, “i”) and are also exactly the same as nominal Instrumental, which is very confusing for learners and would require a lot of cognitive effort to differentiate. Therefore, these are not practiced until Session 53 in Semester 2. Also, usually, Plural adjectival forms are introduced after Singular adjectival suffixes, but adjectives in Genitive Plural (Session 84) normally do not differentiate

between genders (unlike in Genitive Singular) and that is why they are taught before adjectives in Genitive Singular (Session 86).

The sequence of verbal forms is also carefully structured considering learners' processing and production needs, but it is not discussed here, due to the scope of the present study and its focus on case inflection. The reason why verb inflection is listed in the full syllabus in the Appendix L, is to indicate the amount of cognitive processing required from learners in a particular lesson. Care was taken not to introduce new case forms together with new verbal inflection, to allow enough cognitive resources to be directed to encoding and retrieval of new inflection information.

It might be worth noting that the proposed syllabus does not reflect the sequencing of input and output activities, nor interactive practice, which are proposed in 4.2.2 and should be a compulsory element for each lesson within the course. However, as it can be seen from the Appendix L, each grammar topic is closely tied up to a suitable speaking topic which would require that grammar, at the same time, providing opportunities for relevant grammar input and output. For example, Acc. directionality context is matched with "Going places" topic (Session 13), and Genitive Singular after the preposition *dlâ* (for "for") is introduced for "Christmas Shopping" topic (Session 37), at the same time as Instrumental (initially introduced in Sessions 24 and 25) is revisited and is now used in greetings for special occasions (see sessions 48 and 49). "Initial" Dative suffixes are paired with "Visiting friends and family topic.

With regard to testing, I believe that learners' grammar needs to be tested in at least three different modes, at least at the beginner level – in traditional grammar tests (assessing learners' declarative knowledge), in written production (which is likely to assess conscious application of that declarative knowledge) and, finally, in oral production (assessing grammar in use which is edging towards proceduralized knowledge**).** These are the three levels of learning, which the proposed framework aims to help learners to process grammar forms at, as explicit explanation help them form declarative knowledge, then activities for practicing Constituent Assembly and Function Assignment in new contexts train them to apply the newly acquired knowledge to new contexts, and, finally, oral production and interactive activities set the conditions for proceduralization of the grammar form production (see 4.2.2 for discussion). In addition, as a lot of standard language testing is conducted in written form, I believe that it is important for learners to have the experience of written grammar tests, which is likely to be crucial at more proficient levels of language learning. (The problem of testing grammar in oral production is discussed in 5.4 and will not be debated here.) In the

proposed syllabus, two Traditional Written Grammar Tests are scheduled for Session 40 and Session 91, with the assumption that learners would be writing a standard Grammar Exam at the end of the Year. Two essays (scheduled for the Christmas break and the last week of Semester 2) can be part of Continuous Assessment. Six Oral Production Tests can be formative or summative and are distributed throughout the year, but there are more of them in Semester 1, in order to monitor the initial development of learners' grammar system, when a variety of inflections are introduced. The full list of proposed tests within the new syllabus in presented in Table 8.

Table 8

Tests Proposed Within the New Syllabus

Week of testing	*Type of Test*	*What is tested*
Week 2	**Speaking Test 1** (see details in 5.4.3) (can be recorded and emailed to the instructor; can be formative or summative)	Accuracy of oral production of Prepositional suffix "e" as opposed to Nominative (base form)
Week 5	**Speaking Test 2** (see details in 5.4.4) (can be recorded and emailed to the instructor; can be formative or summative)	Accuracy of oral production of Prepositional vs Accusative (suffixes and contexts studied by this point)
Week 7	**In-class 5-min. presentation** (can be prepared during the Reading week, if the University has one; is normally summative)	Accuracy of oral production of Instrumental and Acc. (in Object function)
Week 9	Traditional **Written Grammar Test 1** on the case forms studied (summative)	Metalinguistic knowledge of full set of nominal suffixes for Prep., Acc. inanimate, Instrumental and Genitive Singular, including 7-letter spelling rule
Week 11	Loosely-structured **Interactive Test** - conversation with a peer "Plans for Holidays", 8-10 minutes, (can be formative or summative; if formative, can be recorded and emailed to the teacher)	Accuracy of oral production of all the case forms which were tested in the last written Grammar Test + animate Acc. and Future Tense
to be submitted by Week 12	The first **Essay** 150 words, (normally summative)	Accuracy of written production of the same case forms as above, plus the three tenses
Week 14	**Speaking Test 3** - a simple video episode can be summarized in 7-10 sentences (can be formative or summative)	Accuracy of oral production of Dative, some adjectival case

		suffixes studied and the three tenses
Week 18	**In-class presentations** about different cities in Russian-speaking countries, include unrehearsed questions from class and teacher (can be formative or summative)	Accuracy of oral production of full set for Genitive, as well as some adjectival (not Genitive)
Week 21	Traditional **Written Grammar Test 2** (summative)	Metalinguistic knowledge of Dative, Genitive (including adjectives) and verb forms
Week 22	The second **Essay** on any of the topics studied in Semester 2	Accuracy of written production of all case forms studied in Semester 2 – Dative, Genitive, incl. adjectival

I believe that this kind of combination of different kinds of grammar tests would enable learners to "internalize" the case inflection that they study during the year and successfully use it in oral and written production, considerably more accurately, thus reducing large variability in the summative test results.

In conclusion, I believe that the kind of structuring of learning material, which is described above, would benefit beginner learners immensely, as it would allow not to only learn *about* case suffixes (which is declarative) but use them in their speech (which is working towards procedural). At the same time, learners would acquire the understanding of how the whole inflection system works, enabling them to generalize what they have learnt to new contexts (and therefore apply it to their other learning). The significance of this is wisely reflected in Bruner's famous quote "Knowing how something is put together is worth a thousand facts about it" (Bruner, 2009/1960: 183), and emphasises the main difference between the proposed framework and CLT – the spiralling of grammar is systematic, rather than being needs-specific. So, instead of filling the gaps in learners' knowledge, my proposed framework aims to provide more of a linguistically-informed, conceptually-based approach to teaching, which would be more reliably consistent in helping learners build their knowledge, both declarative and procedural, introducing some deliberate measures to optimize processing and to facilitate proceduralization in instructional settings. I believe that all of these would contribute to increasing the grammar accuracy in learners' production and, consequently, decrease the variability between learners, allowing more learners to achieve higher standards in their language studies.

Among the aims of the present study, there is empirical testing of the effectiveness of the proposed framework application described above. The first section of the proposed syllabus (which covers 10 initial sessions) is used in my longitudinal teaching intervention, whose design is presented in Chapter 5. My previous research project on spiralling, as well as the piloting of the learning materials and the testing tools for the current intervention are described in the next section.

4.2.5. Piloting of the spiralling framework, learning materials and test tools

Having successfully used my spiralling materials with several Russian beginners' courses within the UK Adult Education sector, including one at Swansea University, I have also tested the effectiveness of SC application to the teaching of Russian pronunciation simultaneously with teaching to read in Cyrillic, in the intervention conducted during my MA.

The study aimed to measure the differences in learners' performance, with regards to stress production, vowel reduction and palatalization in Russian words, resulting from different types of instruction. Originally, the research design included two experimental groups and a control group. In the absence of the control group (due to very low response), in order to increase the validity of the results, I recruited the third group, which was taught employing spiralling, but under different conditions.

All participants were native English speakers and complete beginners with regard to Russian (confirmed by the preliminary questionnaire): two groups from Russian Studies (RS) Year 1 and one from other University departments (Group 3). All three groups were taught for 8 weeks, with weekly one-hour sessions, focusing on the same pronunciation skills. The linear group was taught using a standard Russian beginner textbook, different from the one used for their University course. The two SC groups were taught using my own materials but differed in the amount and frequency of practice, as Group 3 had *no* practice outside the experiment, in contrast with the RS groups. Also, in Group 3, pronunciation instruction was integrated with interactive activities.

At the end of the instruction period, all three groups were given the same pronunciation test. The RS groups were also tested for skills retention. The statistical results showed that the two SC groups, produced higher scores than the linear group; with Group 3 outscoring both RS groups, though the difference was significant only with linear group. SC participants also retained the acquired skills significantly better than the linear group. The overall results of this pilot study were very positive, suggesting that the spiralling can effectively facilitate the

acquisition of learners' language skills with regard to Russian pronunciation. I believe that encoding explicit rules for Russian pronunciation and then retrieving them to produce correct pronunciation features in speech, involved processes similar to producing case morphology in speech following the explicit grammar rules. That is why I consider the above intervention very relevant to the present study.

Furthermore, the learning materials from the proposed syllabus for teaching Russian case inflection were piloted in the second half of Semester 2 of Year 1 of my PhD (see Chapter 5 for full methodological details). The time of the piloting was dictated by the necessity to start the main investigations at the very start of Semester 1 of Year 2, due to the designed length (see 5.3.1). Ethical Approval was secured prior to the start of the piloting. Participants were recruited from Russian Studies Year 1 cohort at the home University; they were offered 5 sessions to study Russian cases and practice them in interaction with myself, who is a native speaker of Russian, free of charge, in exchange for taking part in an oral production test. As the purpose of this small piloting was to test run the learning materials and the testing tools (see 5.4), rather than collecting statistical results, the five participants who attended all five sessions, was considered a sufficient number, and the length of the piloting was acknowledged as appropriate. The evaluation of both the materials and the testing tools was carried out and a number of adjustments were made, which resulted in a positive improvement, which ensured the alignment of instruction and testing, as wells as valid testing procedures.

4.3. Conclusion

Having examined the functions and the forms of Russian nominal case inflection from the point of view of learners' cognitive processing, using the two models, discussed in Chapter 3 (WM model and speech production model), I identified a number of characteristics that could potentially cause difficulties for L2 learners of Russian. My brief analysis of Russian beginners' textbooks showed what opportunities are provided to learners to practice different stages of inflection processing, especially encoding and multiple retrieval, as well as the gaps which need to be addressed with regard to my inferences in 3.4. In the second half of this chapter, I outlined the main principles of my new teaching framework and showed how it would restructure grammar material within each session and within the entire beginners' curriculum, in order to enhance learners' processing and address those inferences drawn from research in Chapters 2 and 3. Using examples from Russian case inflection, I demonstrated that using spiralling as the basis and following its three main principles, would allow learners to effectively encode the explicit information about inflection, as well as having opportunities to 'notice' and process case suffixes and functions in the input, and then retrieve them in various contexts during oral production, including unsupervised interaction, when they would have to execute

both Function Assignment and Constituent Assembly. The main research question of the present study is to investigate the effectiveness of the proposed framework, and my research design and methods will be described in the next chapter.

Chapter 5
RESEARCH DESIGN AND METHODS

This chapter sets out the study's Research Questions, which have been derived from the research literature analysis carried out in Chapters 2 and 3, as well as from the discussion of some issues of cognitive processing of Russian case inflection and possible applications of my proposed teaching framework (Chapter 4), and which are at the basis of the research design of this study. I also summarize and clarify the reasoning and assumptions that underpin the Research Questions. I, then, provide detailed information about the design methodology, participant recruitment, test instruments chosen and data collection procedures.

5.1. Research questions and the overall design of the study

5.1.1. Research questions

The main overarching question of this study is whether the proposed framework is an effective approach to teaching ab-initio learners, who have no previous knowledge of Russian or any linguistic background, for them to speak Russian with high inflection accuracy to the expected level of acquisition (as explained below). To answer this question, the study investigates participants' inflection accuracy in oral production tasks and some of the factors that could potentially influence case inflection accuracy in speech (these are discussed further). Other potentially confounding factors, such as age of participants, language background, previous exposure to Russian, among others, were rigorously controlled (see 5.2.1 for details), so they could not intervene with learning gains.

For the purpose of this study, the large and rather general overarching question above was split into more specific Research Questions.

To start with, out of six Russian cases (see 4.1.1), I have selected two, namely Acc. and Prep. (see 4.1.1. and 4.2.2). The choice of cases was determined by several reasons. First, in Chapter 4, these were identified as the most suitable for learners' processing at the start of case-learning. Second, they appear to suit well the structures, contexts and the limited vocabulary available to beginners at the very start of learning, thus are easier to integrate into learners' production. Third, according to Cohine et al's (2019) study, investigating Russian Learner Corpus, Acc. and Prep. account for the overwhelming majority of errors in prepositional phrases, 70.48% (Cohine et al., 2019), which is indicative of acquisition

difficulties. Furthermore, in the same study, Acc. is identified as the case which is most frequently substituted by Nominative (Cohine et al., 2019); that implying that the functions of this case might be challenging for learners to establish during the Function Assignment stage of production. Therefore, investigating learners' processing difficulties in the production of these two cases might shed light onto these issues. Fourth, the two cases appear to be rather different from the point of view of learners' processing (see 4.1.2), thus could potentially produce different acquisition trajectories. Fifth, in most typological classifications, Acc. and Prep. belong to separate classes, for example, (respectively) direct and inherent (Babby, 1980), structural (Lasnik, 2008) and lexical (Legate, 2012), pure and prepositional (Durst-Andersen & Lorentzen, 2017), direct and oblique (e.g., Pesetski, 2013); thus, it is interesting to see whether this is reflected in differences in production. Sixth, these two cases were identified as first in the order of acquisition, derived from written assignments across different proficiency levels (see discussion in 4.1.4). Finally, Prep. and Acc. are normally listed first in Russian beginners' language syllabi and appear first in the majority of Russian beginner textbooks (see Appendix A and discussion in 4.2.2). This fact makes them ideal from the point of experimental design, as they can be conveniently "cut off" from the interference of other cases. Considering the above, my Research Questions are narrowed down to looking at Acc. and Prep. cases only.

Furthermore, as the focus of the proposed framework is to enable beginner learners to utilise grammar in their unprepared speech more effectively, that is to construct their own sentences in Russian with higher accuracy, it is the oral production of case morphology that is investigated in the present study. From a more conceptual acquisitional perspective, many of the inflection acquisition studies examine only the comprehension of inflected forms (Morton, Yakimova, and Van Patten, 2011; Taraban, 2004; Kempe & MacWhinney, 1998; among others), without testing their production, as the latter appears to be a rather challenging aspect to test, particularly at lower levels. However, the paramount role of output in facilitating grammar acquisition has been recognised in SLA for a number of decades (see discussion in 3.3.3), as according to Wong and Van Patten (2003), "learners become better processors of input because they have to create meaning as part of having to express themselves" (Wong & VanPatten, 2003: 415). That is why, it appears rather important to research inflection production - that is, the utilization of case suffixes in speech – as it has a potential of helping us better understand processes, involved in functional morphology acquisition and factors influencing it. At the same time, oral production is a crucial part of speaking skills, which learners often consider the most difficult, as well as the most desirable language skill to acquire (European Commission survey, 2012); therefore the current study

contributes to the growing research in teaching speaking. Thus, my Research Questions focus on case inflection *production*, its accuracy and factors affecting it.

As the current study examines case inflection production from the angle of processing difficulties (as discussed in 3.1, 3.3, 4.1.2, 4.2.3 and 4.2.4), the effects of the three factors that have been identified in Chapter 4 as potentially influencing case inflection production, namely, gender, case context and the familiarity of lexis are examined. As suggested in 4.1.2, due to the two-step processing, required for inflecting feminine nouns, they might have lower case inflection accuracy than that for masculine; that is why gender is selected as one of the factors. Then, inflecting familiar items is anticipated to require less processing effort and, consequently, to produce higher accuracy scores than those for new vocabulary. Therefore, familiarity of lexis is the second factor investigated in this study. Furthermore, it is expected that case inflection accuracy in different case contexts (see definition in 4.1.1) might be different depending on a number of factors, for example, gender of nouns in a particular case context or the type of lexical set that the vocabulary used in it, belongs to (see 4.1.4). Thus, the present study investigates how case contexts affect case inflection accuracy.

Finally, as WM is linked to the success of learning (see 3.1), the potential role of cognitive capacity impacting individual learner variation is investigated. Therefore, the Research Questions include probing possible effects of WM on learners' success in producing Russian case inflection.

In view of the above considerations, I posit the following Research Questions (RQs):

1) Does the teaching intervention using the proposed framework ensures the level of case inflection accuracy of the beginner participants' oral production, as expected in SLA and in Higher Education (HE)?

2) Does teaching intervention using the proposed framework ensure the expected level of accuracy of the beginner participants' production of Prep. case inflection in unprepared speech?

3) Does teaching intervention using the proposed framework ensure the expected level of accuracy of the beginner participants' production of Acc. case inflection in unprepared speech?

4) Do beginner participants perform differently with regard to Prep. and Acc. case inflection in their production?

5) How do different factors, such as gender, case contexts and familiarity of lexis, affect the case inflection production success rates?

6) What is the range of individual differences in participants' production and does the case production accuracy depends on participants' WM characteristics?

7) How do participants' perceptions of learning Russian case inflection, as well as some possible strategies for its production in speech, shed light on individual variation in inflection acquisition?

5.1.2. Overall design of the study

As noted by Jerome Bruner, studying what is going on in a human mind is so convoluted that it "deserves all the rich variety of insight that we can bring to the understanding …" (Bruner, 1990; xiii). Having considered the complexity of the processes involved in case inflection production, in order to answer the above listed RQs, I have decided to employ both quantitative and qualitative methods, which allowed me to explore these processes from different perspectives, thus adding strength to the design through triangulation (as recommended by Morse, 2016; Johnson, 2014, among others). From various models offered by the mixed method research (see Hesse-Biber & Johnson, 2015, for a recent discussion), the sequential mixed method design (e.g., Mark, 2015) appeared to be the most suitable, as explained in the next paragraph.

As participants' performance within the proposed framework has been identified as the outcome measure in the current study, the core component (Morse, 2016) was to be quantitative, measuring the accuracy in oral production tasks, along with quantitative measures of WM. That was followed by the supplementary qualitative investigations (Morse, 2016), which allowed me to examine some aspects of inflection production from a different angle, for example, to investigate learners' views on what affects it and how much attention they pay to case inflection, as well as their strategies in producing inflection. Thus, both the quantitative and qualitative components of the study were used to answer the main question of this study, stated in 5.1.1.

As a quantitative core component, I have chosen a teaching intervention, which has enabled me to apply my teaching framework to teaching case inflection in the conditions comparable to those of a standard beginners' language course, to collect my quantitative data, as well as providing the opportunity to carry out further qualitative investigations of participants' perceptions, difficulties and strategies with regard to case inflection that they were learning within the instruction period. In addition, teaching intervention has been suggested as a suitable way to investigate the role of WM in L2 learning in a new, more authentic setting (Juffs & Harrington, 2011).

The intervention involved recruiting, instructing and testing English-speaking volunteers with no previous knowledge of Russian, who were taught a course of Russian from zero level, using the proposed spiralling framework. In this quantitative part of the study, participants' accuracy of production of Prep. and Acc. case inflection in speech was tested in three rounds of speaking tests (see 5.4.3, 5.4.4 and 5.4.5 for details). In addition, to examine psycholinguistic aspects of inflection processing, participants were asked to complete a Working Memory Test (see 5.4.6), to investigate the impact of Working Memory span on individual variation in inflection production.

For the qualitative part of the study, I have chosen three instruments – learner diaries, questionnaires and interviews, in addition to a qualitative analysis of learners' variation in inflection production. All of these aimed to help me better understand the process of the development of inflection production by analysing learners' perceived difficulties in producing case inflection, their strategies in dealing with rich morphology and their individual approaches to choosing a suffix in a particular context, as well as establishing any overlapping patterns with learners success rates. The three tools provided me with the different levels of depth of investigation, as learner diaries were completely unstructured, while semi-structured questionnaires guided participants towards specific issues related to inflection, and finally semi-structured interviews provided me with the opportunity to explore some answers given in the questionnaires and some entries in the learner diaries, as well as some interesting elements of the test data. The structure of these instruments and procedures of qualitative data collection are described in detail in 5.4.

To sum up the above, this mixed method study investigates the application of the proposed framework to teaching Russian case inflection production in an L2 classroom environment (focussing on the accuracy of Acc. and Prep. case inflection). At the same time, it examines some factors affecting learners' processing, as well as individual variation in the case inflection production. The sequencing of different research methods, quantitative and

qualitative, serves as a basis for the quantitatively-driven mixed method study, with details outlined below.

5.2. Participants

5.2.1. Teaching intervention participants

As discussed in 3.2., the onset of learning has considerable potential for proceduralization to occur and, therefore, investigating inflection production at beginner level appears to be the most beneficial. Also, I expected that the proposed teaching approach should work most effectively if applied from the very start of learning, as it has a potential to help learners build their L2 grammar system, which would be at the basis of their interlanguages, more systematically. This logically narrows the scope of research to beginner level learners, making the study manageable within the time allocated. Moreover, beginners are considered to be the largest group of language learners (for example, in 2013, Introductory to Advanced language courses ratio for the USA Higher Education was 5:1 (Goldberg et al., 2013)). That is why applying the new instructional approach to beginner level and examining its effectiveness would make the results of the current study generalizable to the largest language learner population, which is ab-initio students and non-language degree beginner learners within HE, as well as to non-academic learners, and consequently having the potential to contribute to the development of teaching Russian grammar within HE and beyond. In addition, the absence of prior knowledge of Russian makes participant groups considerably more homogenous by eliminating any interference from previous teaching and learning.

For the teaching intervention, I recruited volunteer learners from other university departments and planned to use Russian Studies (RS) participants as a baseline comparator group (more details in 5.2.2). This meant that, while RS degree students would, for obvious reasons, represent the maximal end of the scale, with regard to L2 input, time investment and motivation for language learning, volunteer participants would be at the opposite end, with limited input within a restricted instruction period, with not much time to devote to their Russian, and with rather low motivation for grammar learning.

Using convenience sampling, I recruited volunteers from university students of various non-language departments. As an incentive, the course of tuition in Russian was offered free of charge, but no extra credits were included in the offer. As the home university did not seem to have a universal system to reach out to all undergraduates, and the Students Union had a

policy of not advertising calls for volunteers, recruitment proved to be quite a challenge, as each school had to be approached individually and not all schools appeared to be supportive. To attract additional participants, approaching neighbouring institutions was considered; however, could have affected the homogeneity of the RS participants, as the syllabus and teaching methods could potentially be different. Ethical Approval for this project following full University protocols, was gained 7 months prior to the teaching experiment, which allowed time for piloting (see 4.2.4). The Information Sheets (Appendix B) were emailed to those who expressed their interest, straight after they responded to the call for participants.

As noted above, getting sufficient participants was a risk for the study. The number of subjects in research literature on inflection acquisition appears to vary between 96 students (Bowles, 2004) and 22 learners of Russian (Kempe & MacWhinney, 1998), which has still been acknowledged as sufficient for viable acquisition testing, as the latter study is well-recognised and extensively cited. Considering the length of the intervention (see 5.3.1.), there was a danger of some participants withdrawing (or being withdrawn) from the experiment, with the number of learners completing the instruction period, being lower. To allow for some drop out, the goal was to recruit between 40 to 50 participants, yielding three groups of 13 -17 learners, which is a reasonable working number for a language class, with the view of possibly merging the groups in Semester 2, if the numbers decreased substantially. This proved to be a viable strategy, though groups were not merged due to scheduling reasons. (In each group, there was at least one participant who could not switch to a different day. Thus merging the groups would have meant losing some of the participants.)

To attend to the homogeneity of the participants from the point of view of language acquisition, the recruitment was restricted to monolingual English speakers who were born in an English-speaking country. Any other language background (including bilinguals) had a potential to skew the results, as it is commonly acknowledged in SLA research that speakers of different languages differ in the way they acquire the same linguistic features in L2, due to L1 positive and negative transfer (Montrul, 2010; Ellis, 2006; Schwartz & Sprouse, 1996; Gass & Selinker, 1992; Odlin, 1989; among others). As discussed above, the participants had to be complete beginners, that is, those who did not know the Cyrillic alphabet and had not had any reasonable exposure to Russian prior to the start of the instruction period. So, heritage speakers of Russian (that is, those who were exposed to Russian within a family environment from birth or an early age) (see Giancaspro et al., 2015; Polinski, 2015; among

others) were not considered. Similarly, so-called “false beginners”, who might have taught themselves Russian or attended a Russian course in the past but considered their language knowledge inadequate, were also excluded. Furthermore, participants were deliberately recruited from non-language schools and departments, to eliminate the influence of previous metalinguistic knowledge of language structures as well as the possibility of established cognitive strategies in language learning.

To ensure that the participants matched the above criteria, a short (one A4 page) structured Preliminary Questionnaire (see Appendix C) was emailed to all 91 respondents immediately after their expression of interest was received. As a result of applying the above filters, 45 teaching intervention participants were selected at the start of Semester 1. The reason for a considerable number of the respondents being screened out was the difference in understanding, by those who were interested in taking part in the project, of the term *native English speaker*, used in the call for participants. Quite a number of respondents interpreted it as a requirement of fluency in English. In the future, I would give a more detailed description of this criterion in the initial promotional email. The same questionnaire was used to rectuit eight RS participants (see 5.2.2). Finally, Preliminary Questionnaires supplied respondents’ age and the department where the respondents studied or worked.

The participants who were selected for the teaching intervention, were sent an email informing them about the selection and inviting them to attend an Introductory Meeting, where they were able to ask questions and discuss various issues regarding this project. If they agreed to take part, they were asked to sign the Consent form (Appendix D). These documents, together with the hard copies of group lists, used to register the attendance, were securely stored in a locked up drawer. On completion of the present study, the hard copies of attendance lists will be shredded. The digital copies of personal information will be kept until the degree is awarded to the researcher, then wiped of the drives soon after the award ceremony. For anonymization purposes, each participant was assigned a number and all their data are stored under this number. A list of these numbers and matching names is stored only in digital form on the University’s M-drive, and will be destroyed together with attendance lists after the degree award ceremony.

Of the 45 participants, who were accepted for the teaching intervention, 7 were non-starters for logistical reasons - the majority of them could not attend the learning sessions at the times which were available; thus bringing the number of participants starting the instruction period to 38. By the beginning of Semester 2, four participants had withdrawn due to the increasing demands of their degree course (e.g., a field trip, time deficit, timetable changes);

three more dropped out due to family circumstances (e.g., moving house, divorce); and three were withdrawn, because of poor attendance or missing several catch-ups (see 5.3.1.). Considering that participants did not receive any credits for taking part and the course was not part of their degree, this rate of drop out in Semester 1 was expected. In Semester 2, on the contrary, the participant retention rate was excellent, as out of 28 learners, who started Semester 2, only one participant could not transfer to online learning after the start of COVID-19 lockdown (see 5.3.1), due to the absence of Internet access. Thus, 27 participants completed the experiment and all the tests required (except delayed testing, see 7.4.). Serendipitously, the same number of A1 level participants are reported by Cherepovskaia et al (2021).

The final 27 participants were relatively even in respect to age (with one outlier) - a mean age of 23, with most ranging from 19 to 28 and one outlier aged 35. The male-female ratio was slightly in favour of male participants (15 in comparison to 12 female learners), similar to numbers reported in some studies listed above. Participants represented the following University departments: Mechanical Engineering, Geography, Computing, Astrophysics, History, Mathematics, Earth and Environment, Chemistry, Medicine, and Electronic and Electrical Engineering. This representation was not controlled, apart from not including language departments, as mentioned above.

All recruited participants were assigned to a learning group, depending on which day of the week they could attend learning sessions. Best efforts were made to make groups as homogeneous as possible, for example, the number of male and female participants was similar in all groups. It is possibly worth noting though, that a perfect split of participants between groups in the real-life teaching scenarios, which the current study aims to replicate, is hardly possible. In Semester 2, the groups were not merged, as explained above, but the number of learners in each group was kept balanced - no less than 9 in a group.

5.2.2. Russian Studies participants

The original plan was that these RS students would constitute an appropriate comparator group, which would provide a reliable guideline for a level of case production that needed to be reached by beginner volunteers (rather than a native speaker control group). A call for participants was sent via a department email channel, and five hours of speaking practice with a native Russian speaker, that is me, free of charge, was offered as an incentive for participating, gaining this study 41% of the Year 1 RS population at the home University.

However, only eight participants were selected with the help of the Preliminary Questionnaire, excluding bilinguals, heritage speakers and prolonged “false beginners”, that is those who consider themselves beginners, but might have had some exposure to Russian prior to the start of learning, for example, taught themselves some Russian, lived in a Russian speaking country for an extended period of time, or had a Russian relative whom they communicated with.

I considered adding more numbers to this group, which I would have had to source from other institutions, which would have reduced the homogeneity of the group and would have introduced even more confounding factors. The baseline group was homogeneous with regard to age, as all of them were 18 and 19 years old, but opposite in respect to gender, as there was only one male among 7 females. That could not possibly have been addressed as the recruitment depended on limited number of respondents from a small pool of Year 1, which consisted predominantly of female students (61%).

The RS participants completed their oral production test and attended their five speaking sessions, but I was still concerned with low numbers, as well as a number of confounding factors. As a solution, I decided to assess the outcome measure of accuracy in the present study against the accepted level of feature acquisition (of 60%) in SLA (e.g., Vainikka & Young-Scholten, 1994) and the standard HE grading percentages instead (thus not employing these data in my analysis).

5.2.3. Second raters

The last kind of participants were Russian speakers recruited for second rating of the language tests, used to increase validity and reliability, by addressing any risk of researcher’s subjectivity in the analysis of the tests. Due to the very low language level tested, raters were not required to have any linguistic or teaching background. Any native Russian speaker educated in Russia to a reasonable level (e.g., allowing them to enter a UK University), was suitable to perform the task of marking the transcribed tests according to the marking scheme provided (see 6.1.2.). The raters were invited via Student Union channels. Two undergraduate students from Russia, possessing the above standard, were recruited from the home University. In return, they were offered some sessions (depending on their time investment into the test marking) of tuition in English as L2 by me, free of charge. The raters’ age or gender are of no significance to this study, as these would not affect their grammar knowledge, required for the rating.

5.3. Teaching intervention

5.3.1. Instruction period and transition to online during the COVID lockdown

The research studies on Russian inflection acquisition, discussed in 4.1.4, vary in respect to the length of the experiments, from 10-15 minutes of experimental instruction within a standard beginners course sessions (Arnett & Lysinger, 2013) to 6 training sessions within 14 days (Kempe and Brooks, 2008). As far as my investigations could stretch, the longest teaching intervention reported in recent research literature appears to be a 12 hour (over 12 weeks) experiment by Derwing and Rossiter (2003), which tested L2 English pronunciation and was run as a course in pronunciation for University language degree students of intermediate level.

The teaching intervention for the current study was designed as a 20 hour experiment for several reasons. With participants being complete beginners (see 5.2.1.), the length of the instruction period was dictated by the necessity to prepare, within the first half of the intervention, the language database required for the teaching of cases. For example, in order to process meaningful input containing required cases, learners must be able to comfortably read words written in Cyrillic and acquire some very basic vocabulary. Moreover, in order to produce inflected case forms in varied contexts and to engage in speaking activities, learners needed to understand a category of grammatical gender and have the idea of verb conjugation. As my previous teaching intervention (see Parker, under review) demonstrated, with the use of spiralling, English speakers are able to acquire solid reading-in-Cyrillic skills and master a reasonable arsenal of simple words and sentence structures, as well as basic concepts of parts of speech, within 9-10 hours; thus, the first half of the intervention was planned for 10 hours.

The second half of the instruction period needed to be long enough for learners to master the “initial” case suffixes and some of the first contexts for the *two* selected cases (see 5.1.1) in their speech. As demonstrated in my proposed syllabus, discussed in 4.2.4 and presented in Appendix L, it is possible to do within ten weeks. From my extensive teaching experience, including teaching Year 1 Russian language course at the home University during Year 1 and Year 3 of my PhD, 5 hours appeared a feasible length of instruction for one case (also see Rifkin, 2005), thus the second half of the intervention had to allow for at least 10 hours.

Considering the above, the teaching intervention was designed as a 20 hour Russian beginner course, in which the content was similar to that of a standard beginners' Russian language class, but was restructured according to the proposed grammar teaching framework (see 4.2). As participants were non-language degree students, it seemed sensible to offer them weekly sessions, which they were more likely to be able to fit into their timetable, rather than a more condensed course. From a logistics point of view, the total of 20 weekly sessions fitted well into two University semesters, when participants were easily available. On the other hand, a two-semester course required a bigger commitment on their behalf, which had always had a possibility of substantial drop out (see discussion in 5.2.1), as well as the potential for some unforeseen changes in circumstances.

One of such changes was the start of COVID-19 pandemic, when during Week 17 the teaching unexpectedly had to be transferred online, which required additional effort and swift action from me to retain the participants. In order to ensure that there were no gaps in the Instruction period (which potentially could have affected participants' learning) and keep my participants engaged, I had to make all necessary arrangements over the weekend for the online learning sessions to start online the following week.

First, I had to find a suitable online platform. Skype was chosen over other online platforms mainly for the three following reasons: 1) ease of transition (Microsoft Teams, suggested by the home University, required time for authorisation for setting up groups and was unfamiliar; Zoom was not recommended by the home University due to controversial security issues that it had at the time, and required separate consents signed by the participants, which would have delayed the transition); 2) familiarity and ease of use for the teacher, which I believe is very important for interactive online speaking sessions (I have taught freelance using Skype for the previous 5 years); 3) pedagogical convenience, as teacher's notes in Skype Chat are retained for future use by students (unlike in Zoom, where the meeting chats are not saved).

Then, I had to act quickly to make sure that all 28 participants were informed and had a chance to ask me questions if they had any. I communicated via University email and encouraged all participants to come back to me, should they have any difficulties or if they are unsure or need clarifications. The majority of invitations to Skype were emailed to participants individually within a couple of days from the University being closed for the lockdown. A detailed explanation on how to join Skype was created and emailed together

with each invitation. The online sessions were scheduled on exactly the same days and at the same times, which I believe was crucial for my participants to stay with the project. The above arrangements enabled the last four sessions (Week 17 - Week 20) to be successfully completed online. As a result, there were no gaps in the second half of the intervention, caused by the lockdown, and out of 28 participants, who started Part 2 in February, 27 completed the experiment.

All experimental groups received full course of 20 hours of tuition in Russian over two semesters with weekly sessions lasting one hour. The learning sessions were scheduled for the same time (4 – 5 pm) on different days (Tuesday, Wednesday, Thursday), attempting to keep learners' processing conditions as similar as possible.

Participants who were offered a place on the project and agreed to take part, were asked to attend all 20 scheduled sessions (unless they decide to withdraw) and treat these as their weekly commitment for the duration of the two semesters. To ensure that the processing conditions during a session were similar for all participants, measures were put in place for any absentees to catch up with the group before they re-joined. If a participant missed a session, they were emailed the handouts that they have missed, asked to go over them in their own time and offered a 15 minutes catch-up session with myself, *before* the following group session. This information was included in the Participants Information Sheet and reinforced during the Introductory Meeting. All 27 participants, who were tested, either attended all the learning sessions or had all their catch-ups; with the majority having no more than three catch-up sessions in the second half of the experiment (with only two participants having four - one due to illness and the other due to lockdown changes).

5.3.2. Content and materials used

It is important to note that the lesson time, even in the second part of the instruction period, was not solely devoted to practicing cases – participants were reading, learning to speak and communicate, and using cases was part of that process. This is clearly demonstrated in 4.2.3, where the 10 weeks of initial case-learning is explained and illustrated with figures 12a-12f (these are also weeks 5-15 in the proposed syllabus in Appendix L). Along with Prep. and Acc. cases (selected in 5.1.1), during the last 10 hours, new vocabulary and new verb forms were introduced, to enable learners to build a functioning interlanguage, as well as to compare the production of verbal and nominal inflection. The list of grammar topics,

which were covered, and vocabulary pools, which were studied and utilized in speech, are summarized in Table 9 and Table 10.

All three groups were taught using the proposed spiralling framework. During Part 1 of the intervention, experimental groups were taught using materials which have served as the basis for my Russian beginners' textbook "Russian in Plain English: A Very Basic Russian Starter for Complete Beginners", published by Routledge in July 2020, that is, after the instruction period had finished. The materials have previously been trialled with several groups of adult learners, including Russian beginners' course at Swansea University and during my first teaching intervention for my MA (see 4.2.4) (samples are available). The materials used in Part 2 of the current intervention, dedicated to teaching cases, had been piloted during the first year of my PhD (see 4.2.5) and were evaluated and developed further.

All of the materials strictly adhered to the principles of the proposed spiralling framework (see 4.2); that is, the explicit information was segmented and initially presented in small amounts, as well as basic case concept being explained in a simplified form; then target forms were included into texts (sentences initially) for reading or small conversations, with 'noticing' encouraged; after that the same forms (*not* the same words) were used with various vocabulary during speaking activities, including interactive tasks; and finally, they were revisited with more information added and new contexts introduced. See a sample handout for Lesson 2-1 in Appendix E.

It feels appropriate to mention that the spiralling employs a non-paradigmatic approach to grammar, which implies that grammar classes are not taught as a whole and not necessarily in a traditional order, with the main purpose of enabling learners to process the explicit knowledge effectively and use it in practical language skills, here speaking, through more effective encoding and multiple retrieval. For example, first-person pronouns and third-person pronouns are taught first as they do not require turn-taking, then second-person Plural is added, as it enables learners to ask and answer questions, using polite forms. No Plural case forms were taught, firstly, in order to make the scope of the study manageable; secondly, to comply with spiralling principle of introducing the information in small amounts; and thirdly, to comply with the majority of beginners' syllabi, according to which Plural is taught after Singular case forms. Table 9 presents the lists of grammar classes, inflected forms and some other grammar features which were introduce to participants during the Instruction period.

Table 9

Grammar Topics Covered Within the Teaching Intervention

Part 1	Week 1-10	Gender	Personal pronouns (1st and 3rd person Singular) Nominal gender Possessive pronouns and gender agreement Negative sentences Flexible word order Personal pronouns (2nd person) "I/you have" construction Nominative case (Subject function) Verb conjugation (Infinitive; 1st person, Singular)
Part 2	Week 11-15	Prep.	Prepositions *v* (for "in", "at", "to") and *na* (for "on", "at", "to") Verb conjugation (2nd person Plural & 3rd person Singular)
	Unprepared Oral Production TEST 1		
	Week 16-20	Acc.	Possessive pronouns (3rd person Singular) "he/she has" construction Adjectives
	Unprepared Oral Production Test 2		

The choice of vocabulary pools (see Table 10) complies with the proposed syllabus in Appendix L, and includes case contexts allocated for Prep. and Acc. for the first 10 weeks of case-learning. This vocabulary enables learners to use the case suffixes which were introduced, in their speech. Finally, this vocabulary meets the requirements of the beginners' syllabus, thus this study could be considered generalizable to the ab-initio student population.

Table 10

Vocabulary Pools Introduced within the Teaching Intervention

Part 1 (Semester 1)	**Part 2** (Semester 2)
Nouns: Things People	*Nouns*: Buildings Sports games

Places	Musical instruments
Foods	Days of the week
Verbs (stationary, e.g., *rabotat'* for "to work"; only in 1st person Singular, enabling learners to say *Â ne rabotaû. Â student.* for "I do not work. I am a student." *or Â student, no â rabotaû inogda.* for "I am a student, but I work sometimes", not requiring cases.)	Months
	Transport
Adjectives (for languages, e.g., *russkij* for "Russian")	*Verbs* (stationary, e.g., *žit'* for "to live", requiring Prep.; *igrat'* for "to play", which can agree with Prep. or Acc., depending on the context;
Adverbs (e.g., *horošo* for "well", *ploho* for "bad/ly")	two verbs of motion, requiring Acc.: *idti* for "to go on foot" and *ehat'* for "to go by transport")
Demonstrative pronouns (e.g., *tam* for "there", *èto* for "this")	

Though gender acquisition is not the aim of this study, gender was identified as one of the factors that could influence the accuracy of case inflection production (see 4.1.2). That is why the best effort was made to keep a reasonable balance between the use of feminine and masculine nouns in the instruction materials, as well as in the test materials (see 5.4.3., 5.4.4. and 5.4.5.). Similarly, as case contexts are another factor that is investigated, I have ensured that, for each of the two cases, nouns representing the required case contexts, as well as all three lexical sets, were present in participants' speech – closed, limited and open (see 4.1.2).

Following the principles of the proposed framework the case information was segmented and spaced throughout the Instruction period – Prep. was introduced first and Acc. was started five weeks later. During Part 2 of the Instruction period, case contexts for Prep. were alternated with those for Acc.. The order of introduction of vocabulary pools, and consequently the case contexts is clearly demonstrated in 4.2.3.

It is important to note that each suffix at each stage of instruction was recycled in very simple interactive activities offered for speaking practice, in between explicit explanations and reading, in order to address the inferences made in 3.4. Although this study was investigating the production of case suffixes in speaking, reading was an essential part of the instruction for a number of reasons. First, in the classroom environment the majority of the L2 input is normally provided through reading. Second, reading was used for pushed 'noticing' during the processing of inflection in the input (recognition) (see 3.3.2), as an important facilitator of inflection encoding. Third, visual comprehension of inflected forms in a text would appeal to learners who have a sensory preference for visual processing, and who

constitute the majority of learner population (Dörnyei, 2014). Fourth, reading aloud assists learners in developing their Constituent Assembly (See Figure 2 in 3.1) skills, paramount for oral production, as well as establishing case functions for further Function Assignment (see Figure 2 in 3.1); thus, preparing the ground for their recall in speech production. Finally, the content of the instruction for the experimental groups needed to be as close as possible to that of a standard Russian beginners' course, where reading constitutes one of the main activities.

None of the learning activities, aimed specifically at teaching writing or listening, were included in the instruction, due to the focus of the current study on production, in addition to the time constraints. According to typical pedagogic practices, exemplified in recent discussions at relevant conferences (e.g., AATSEEL, American Association of Teachers of Slavic and East-European Languages; and MGIMO, Moscow State Institute of International Relationships), as well as to my extensive personal teaching experience, this is the strategy often employed by Russian language tutors, leaving listening and writing as homework, allowing as much of classroom time as possible for explicit explanations and for oral production, which are difficult to practice at home. Participants were made aware of this condition and were given guidance on how to practice those skills, if they wished.

The learning materials contained all necessary explanations, comments, vocabulary notes, exercises and activities required. During the instruction period, for each session, each participant was provided with a photocopied set of handouts in A4 size, which they were allowed to keep. After transferring to online teaching in Week 17, the handouts were emailed to participants in PDF format, before each session. In order to keep input and output as controlled as possible, no homework was given, though suggestions were made of what learners could do between the sessions, should they wish to. The amount of individual study was impossible to control.

5.4. Quantitative data collection

5.4.1. What is tested

The absolute majority of data for the quantitative analysis for the present study was collected from the same 27 participants who were recruited as Teaching Intervention Participants (see 5.2.1) and taught during the instruction period described in the previous section (5.3). (The exception is the delayed test where just over 50% of the Teaching Intervention Participants

responded, see 5.4.5). As participants recruited for the present study, needed to be complete beginners, no pre-test was required (see preliminary screening in 5.2.1).

In order to answer the majority of my RQs, I needed to measure participants accuracy of oral production of Prep. and Acc. suffixes that participants were learning during the last 10 weeks of the instruction period (the first 10 hours of case-learning), namely, one suffix “e” for Prep. and two suffixes for Acc. - zero suffix and “u” (see Table 6).

As Russian case suffixes can be different for different genders and, in 4.1.2, gender was identified as one of the factors that could potentially affect case inflection accuracy, the decision about which genders to include in the study had to be made. Three aspects were considered – frequency of items of a particular gender in speech, their use in the target case contexts and syncretism of their forms. First, the distribution of genders in Russian is not equal - neuter makes only 13% of nouns in native Russian, compared to 46% masculine and 41% feminine, (according to Muchnik, 1971). Exactly the same percentage (13%) of neuter nouns was quoted from Russian National Corpus (807,491 neuter from a total of 6,102,386 Nom. Sing. Nouns). Thus, I did not expect neuter to deliver significant number of case forms. Second, limited number of neuter nouns that are present in standard beginners’ materials, due to their semantics, are highly unlikely to be used in the case contexts which were taught during the instruction period (for example, an apple is not a likely location). Third, some neuter nouns, which could be used in target case contexts, do not indicate a change in case for Prep. and Acc., either because they are indeclinable (e.g., *kino* for “cinema”), or remain the same in both Prep. and Acc., as their base-form has “e” as a neuter marker and is syncretic to both case forms (e.g., *voskresen'e* for “Sunday”). Thus, it is impossible to tell whether these items are inflected for case or not, from their form. For these reasons, from the three Russian genders, I have decided to concentrate on masculine and feminine case forms and exclude neuter from the data. I believe that case inflection production for neuter would be interesting to be investigated in a separate study, which would carefully address the limitations listed above. However, some neuter items were present in sentences produced by participants, because they were included into learning materials, in order to comply with an appropriate Russian beginners’ curriculum and to answer the likely needs of the learners’ communication context.

Considering the above, the four case suffixes were selected for testing – two for Prep. (masculine and feminine) and two for Acc. These are presented in **Table 12.**

Table 12

Four "Initial" Case Suffixes Which were Taught and Tested in the Present Study

	Masculine	**Feminine**
Prep.	**"e"**	**"e"**
Acc.	**zero suffix**	**"u"**

As Table 12 demonstrates, the Prep. suffix is the same for masculine and feminine, while Acc., has a different suffix for both genders. However, as gender has been identified as one of the factors investigated in this study and feminine case suffixes are expected to involve different processing mechanisms during inflection production (see discussion in 4.1.2), Prep. suffix "e" for masculine and Prep. suffix "e" for feminine will be always referred to as two separate suffixes. Examining how gender affects case inflection accuracy will answer the RQ5.

Testing the accuracy with which participants produce these suffixes in the oral production in a post-test would answer RQ1, RQ2, RQ3 and RQ4. However, in order to investigate the difference in success rates between genders when the case suffix is the same, the interim test (Test 1) was created. It measured the accuracy of case inflection production for the Prep. case only, *before* Acc. was introduced in Week 16 of the Instruction period. This could potentially demonstrate whether the two-step processing during the inflecting of feminine nouns for case, discussed in 4.1.2 and 4.2.2, requires more cognitive resources than masculine case forms, if the accuracy rates are different for Prep. masculine and Prep. feminine. In addition, comparing the case inflection accuracy in Test 1 with the success rates in the next test (after the Acc. is taught) would allow to see whether the drop in the initial Prep. accuracy could be explained by the increase of processing load caused by the introduction of Acc. So, Test 1 would answer RQ2 and partially RQ5. See 5.4.3 for details.

Then, in the post-test at the end of the Instruction period (Test 2), both oblique cases (Prep. and Acc.) would be produced by participants, which would increase the processing load, because the Function Assignment would become considerably more challenging then in Test 1, due to the choice between different case contexts (for the two cases), in addition to the increased number of suffixes (all four suffixes from Table 12). The post-test data would answer the first three RQs. Also, to contribute to answering the RQ4, success rates for the two cases would be compared. These data could potentially show whether there are any differences in Constituent Assembly between Prep., which has the same suffix for both genders, and Acc., which has different suffixes for masculine and feminine (see Table 12). These would be the main aims of Test 2 (see 5.4.4 for details).

Next, to examine how well the skills tested in the post-test, are retained, a test similar to Test 2 is administered to the same participants after a period of time. Therefore the aim of the delayed test, Test 3, is to measure the accuracy rates for the same four case suffixes after 6-8 weeks after the post-test and compare these to those for Test 2. This would contribute to answering the first three RQs.

Furthermore, in order to answer the RQ5, using the same data from Test 2, the case inflection accuracy is also compared between the two genders, namely, masculine and feminine, as well as different case contexts (defined in 4.1.2). The following case contexts were selected for this analysis (see Table 13); these matched those that were taught during the instruction period. (In Test 1 there was only one case context (location) and in Test 3 case contexts were not analysed due to the scope of the study.) As it can be seen from the table below, there is at least one case context for each of the lexical sets for each of the two cases.

Table 13

Distribution of Case Contexts between Different Types of Lexical Sets.

Type of lexical set	*Prep.*	*Acc.*
Closed	Months	days of the week
Limited	musical instruments, transport	games
Open	place names in location context	place names in direction context, direct object

It is also important to note that the relationship between gender and case contexts is far from being straightforward and that gender can be a confounding factor for the case inflection accuracy in particular case contexts. For example, months are all masculine in Russian, while in the days-of-the-week context, the distribution of masculine and feminine items is equal – 3 : 3 (with the seventh day being neuter). If my suggestion about a heavier processing load for feminine inflection (because of the two-step process) is correct, it could be possible to assume that these discrepancies in gender distribution could result in higher success rates for masculine-only case context. In addition, lexical sets can be different for different case contexts (see 4.1.2 for definitions) and can affect accuracy, as much as nouns from the closed lexical sets are likely to be prone to chunking. This means that, within the closed lexical set, for Prep., only masculine case forms will be produced, possibly drastically

increasing the case inflection accuracy, while, for Acc., there will be a combination of both genders, possibly reducing success rates. However, the accuracy for the days-of-the-week context might still be higher than for any other Acc. case context. Within the limited set, the situation is opposite, because, here, inflected items will be all masculine for Acc., as games are predominantly masculine in Russian (with rare exceptions, which are highly unlikely to be found in beginners' vocabulary, e.g., "lapta", an old Russian game which is often explained as a Russian version of baseball). At the same time, musical instruments, which are inflected for Prep., can be masculine or feminine, though, in learners' production, the ratio between genders might not be as equal, as for days of the week, as they are likely to produce the musical instruments that they or their family play, regardless of gender. Finally, in the open set, place names, used in both Prep. and Acc., can be of either gender. The distribution of gender within open and limited sets in learning materials and in test tools was controlled where it was possible (e.g., in open lexical sets).

Finally, in order to investigate the differences in the case inflection accuracy between familiar and new vocabulary (RQ5), in Test 2, some new nouns were included into open lexical sets, namely, unfamiliar place names. The number of new items was equal for both genders, as well as for both cases (see further details for each test tool in 5.4.4 and 5.4.5). The exact definitions for which items were considered familiar and new, are given in 6.1.3 (see Tag 8). Though the familiarity of lexis was not investigated in Test 3, the same number of new items was built into the delayed test, to ensure the comparability of the two tests.

5.4.2. The problem of case inflection testing

In order to demonstrate whether and how accurately participants could apply knowledge of grammar rules, retrieving and producing correct case suffixes and assigning correct case in different case contexts, this study needed appropriate test instruments to capture inflection production in learners' unprepared speech,

To start with, I considered SLA-based tests investigating the implicit knowledge of grammar features, such as, the Grammaticality Judgement Test and the Elicited Imitation Test. Both were deemed not suitable for this study, as the former is based on comprehension, rather than production, while the second, though testing production, relies on listening, which was not taught or practiced by the participants of the current study (see reasoning in 5.3.2). Moreover, I was looking for a way to test inflection production, which could potentially be used in a standard classroom or language exam environment. As far as my investigations

could stretch, neither of the above SLA tests have been as yet widely employed in teaching practice, and seem to require some special training to be implemented (e.g., Erlam, 2008).

Turning then to language testing used in present educational contexts, I discovered that, despite the common agreement that the aim of learning foreign language grammar is its implicit knowledge, that is, its creative use in learners' spontaneous speech (R. Ellis, 2006), traditional grammar testing still relies heavily on written tests (e.g., Larsen-Freeman, 2009), appealing to learners' explicit (declarative) knowledge, which is the formulation of rules. On the other hand, the most common means of testing production skills is a recorded oral exam or Oral Proficiency Test (OPT) in the form of an interview (Fulcher, 2015), which, though effective for assessing speaking abilities, appears to have three issues with regard to grammar. Firstly, it is not normally designed specifically to test grammar; linguistic competence is assessed holistically. Secondly, the first speaking test is commonly administered to beginners at the end of Semester 1, or Semester 2, (after about 50 -100 hours of study, according to Russian major programmes in the US and Russian language modules in the UK), when the majority of basic grammar (including cases) has already been completed; thus, it is unable to offer any feedback for teaching grammar to beginners. Thirdly, Year 1 speaking tests are often based on topics prepared in advance (e.g., Larsen-Freeman, 2009) and on pre-learned chunks, rather than testing the utilization of grammar in new contexts, which does not guarantee testing how well learners could assign functions in new situations or assemble constituents using unfamiliar vocabulary.

This makes it possible to infer that, as a common practice, grammar is taught without tests that aim to assess its production in speech. And yet, it has been suggested that the acquisition of grammatical competence is the best predictor of attainment (Brecht, Davidson, and Ginsberg, 1993) and correlates directly with the acquisition of the four skills (Rifkin, 2005: 12). Thus, it appeared that the problem of how to test the production of grammar features in learners' unprepared speech, in time to address possible issues, that is, at very early stages of learning, was not necessarily restricted to the current study, but, I believe, is underlying grammar testing in general.

Looking for possible solutions for this problem, with regard to testing Russian case inflection, I designed three speaking tests, specially aimed to test the accuracy of case inflection production in learners' unprepared speech.

5.4.3. Guessing Game Test (Speaking Test 1)

The first speaking test was to answer RQ2 and needed to test learners' accuracy of the production of Prep. case inflection ("e" for masculine and "e" for feminine", see Table 12) before participants started learning Acc. (see 5.4.1). Under this condition, Function Assignment should not be taxing for participants' cognitive processing, as location context (Prep.) has to be only differentiated from Subject function requiring Nominative forms, which is normally obvious from a pragmatic context. (In Test 1, Prep. was produced only in location case context). Thus, participants' cognitive resources would be mainly directed at Constituent Assembly, which is expected to be different for masculine and feminine case forms, due to the suggested two-step processing required for feminine (see 4.1.2). Therefore, the aim of Test 1 was twofold – first, to determine whether the accuracy for Prep. would be different for masculine and feminine case inflection; and second, to see whether Test 1 success rates for Prep. would be significantly different from those in the next test, where both oblique cases (Prep. and Acc.) would have to be produced by participants, increasing the processing load.

The first test had to be administered in a very short period of time after the Prep. case had been taught and before the Acc. was introduced, during Week 15 (after Session 5 of the second (case-learning) part of the teaching intervention (see 5.3.1. on instruction period). The test had to be suitable for learner's limited vocabulary and simplistic sentence structures, considering that the first 10 hours of the instruction period participants were focussing on learning to read in Cyrillic (see 5.3.2 on the instruction content), but at the same time, it had to elicit a reasonable amount of unprepared speech with obligatory context for the case.

The designed Speaking Test 1 meets the above requirements and is based on a guessing game, sometimes used in language teaching, though the use of the game for testing or teaching Russian cases has not been cited. The game is played by two learners, who sit on opposite sides of a table with a screen in the middle; each of them having a set of objects, which includes exactly the same items, representing the vocabulary that learners would have learned in the class, for example, a book, a pen, a pencil, a glass, a cup, a bag. The screen has to be tall enough for participants not to see the objects on the table, but low enough to see an object that the opponent chooses to show for the round. The aim of the round is to find where that one object is placed by asking questions in the target language, without using the word for *'Where?'* for at least three goes, e.g., *Ručka v stakan**e**?* for "Is the pen in the glass*[Prep.masc.]*?" or *Ručka na korobk**e**?* for "Is the pen on the box[*Prep.fem.*]?". For the purpose of this test, participants were restricted to using the prepositions *v* for "in*"* and *na* for

*"*on*"*, both requiring the Prep. case. The test was piloted prior to the start of the instruction period with the RS group.

The sets of objects, selected for Test 1, included 14 items, typically used within the first 15 lessons in Russian beginners' textbooks. The selected nouns included six masculine, six feminine and two neuter nouns (see the list in Appendix F). The reasoning for the gender split was discussed in 5.3.2. The two neuter nouns (*âbloko* for "apple" and *moloko* for "milk") belong to beginners' minimum vocabulary (GCSE specification) and were included to provide a variety of nouns to be used in Subject function. I did not plan for these to deliver significant number of Prep. case forms, if any, as they are not commonly used as locations (required to elicit Prep. in Test 1).

Participants took Test 1 "The Guessing Game" face-to-face in Week 15, as it had been planned. The times of the tests were allocated according to participants' availability, which determined pairing up. Each test lasted 6-7 minutes (unless there was a delay or a gap in speaking), allowing between 3 to 5 rounds for each participant, who provided on average 10-17 case forms each in their questions and answers, the majority of which contained both Nominative and Prep. case forms. Most of the participants were getting very engaged and often quite competitive, placing objects in the most unpredictable places, thus supplying a variety of new contexts. The tests were audio-recorded and then transcribed.

5.4.4. Speaking Test 2 (including Comics Test as the second part)

Unlike in Test 1, in the post-test (Test 2), administered at the end of the instruction period, the case inflection accuracy for both Prep. and Acc. was tested. Thus all four case suffixes from Table 12 were in participants' production, considerably increasing their processing load, as Function Assignment had to be executed for all case contexts from Table 13, as well as Constituent Assembly with different overt suffixes for feminine ("e" for Prep. and "u" fro Acc.). Therefore, Speaking Test 2 had a number of purposes:

1) in order to answer RQ1, it tests the overall case inflection accuracy;
2) to contribute further to RQ2, it investigates how the case inflection accuracy for Prep. is different from Test 1, that is, affected by the introduction of the second oblique case, namely, Acc.
3) to answer RQ3 and RQ4, it tests the case inflection accuracy for Acc.;
4) it also provides data for analysing the effect of the three factors which are listed in RQ5 as having the potential to affect the accuracy of case inflection production;

5) to answer RQ6, its data on case inflection accuracy is used to investigate the effect of WM span.

To achieve the above aims, I needed to create a test which would elicit both Prep. and Acc. case forms in participants' speech, including different contexts listed in Table 13.

Initially, I designed an interview test which would elicit the required case forms. The questions were phrased in such a way that they would not provide participants with the target case forms. For example, *Gde ty živëš'?* for "Where do you live?", requiring Prep., would be appropriate, while *Ty živëš' v gorode ili v derevne?* for "Do you live in a city or in a village?", containing Prep., would not. Similarly, *Kuda ty edeš' letom?* for "Where are you going in summer?", requiring Acc., would be selected, instead of *Ty edeš' v Rossiû letom?* for "Are you going to Russia in summer?" that contains this case.

The piloting of the designed interview test with RS students (during Year 1 of my PhD) revealed that the number of case forms that they produced, were quite limited - about 5-7 for the majority of learners). This is also reported by Pienemann and referred to as "poverty of data" (Pienemann & Kessler, 2012, Pienmann & Mackey, 1993). With the purpose to collect more relevant grammatical data, Pienemann & Mackey designed a number of elicitation activities based on task-based learning; these proved to be successful but were aimed at intermediate students. Consequently, I had to find a different way of eliciting unprepared speech from beginner learners. As this study has aimed to investigate learners' underlying difficulties in producing inflection and how they manifest themselves in learners interlanguages (see discussion in Chapter 2), my testing needed to follow Selinker et al.'s proposal to analyse "the speech forms which result from the attempted expression of meaning in a second language" (Selinker et al., 1975: 140), which they differentiate from "language-like behavior" (as defined by Spolsky (1969)). Loschky and Bley-Vroman (1993) noted that it is quite challenging to create a production task that could make production of the target grammatical feature 'essential'. In addition, learners' cognitive processing challenges in language testing needed to be accounted for (see Kormos, 2014, for discussion). The solution was found in the use of comics.

The idea of using comics in language teaching has been around for a number of decades (see Brown, 1977; and Vassilikopoulou et al., 2011, for a review). They have been actively used for developing reading and speaking skills, in L1, as well as L2, particularly within teaching English as a Foreign/Second Language (Krashen, 1993; Williams, 1995; Liu, 2004; Norton & Vanderheyden, 2004; among others), with one interesting study investigating

cognitive effectiveness of comics as an instructional tool (Mallia, 2007). Along with pragmatic orientation, contextualization of learning and engaging content, one of the main advantages, pointed out by many language pedagogy researchers above, is their authenticity. However authentic comics were sometimes criticized for inappropriateness of their contents for learners' age or their language level, as well as for some ethical issues. Recently, there appeared a few papers advocating the creation of comics specifically for developing particular language skills (Graham, 2011; Vassilikopoulou et al., 2011; among others).

Despite the growing interest in the use of comics among the researchers, I was unable to identify any studies where comics were used for teaching morphology, particularly case forms. Neither could I find any research investigating the use of comics in language testing. However, the use of images in teaching grammar was reported at a conference at MGIMO University in Moscow (Maksimenko & Belyakov, 2018) and was based on the work of a Russian school teacher A. Belovitskaya, who teaches L1 literacy using her own pictures, which, though would not, strictly speaking, comply with the definition of comics as a genre, are very popular among pupils and teachers alike (https://www.adme.ru/svoboda-kultura/russkij-yazyk-v-kotah-432855/), and have been published and developed since.

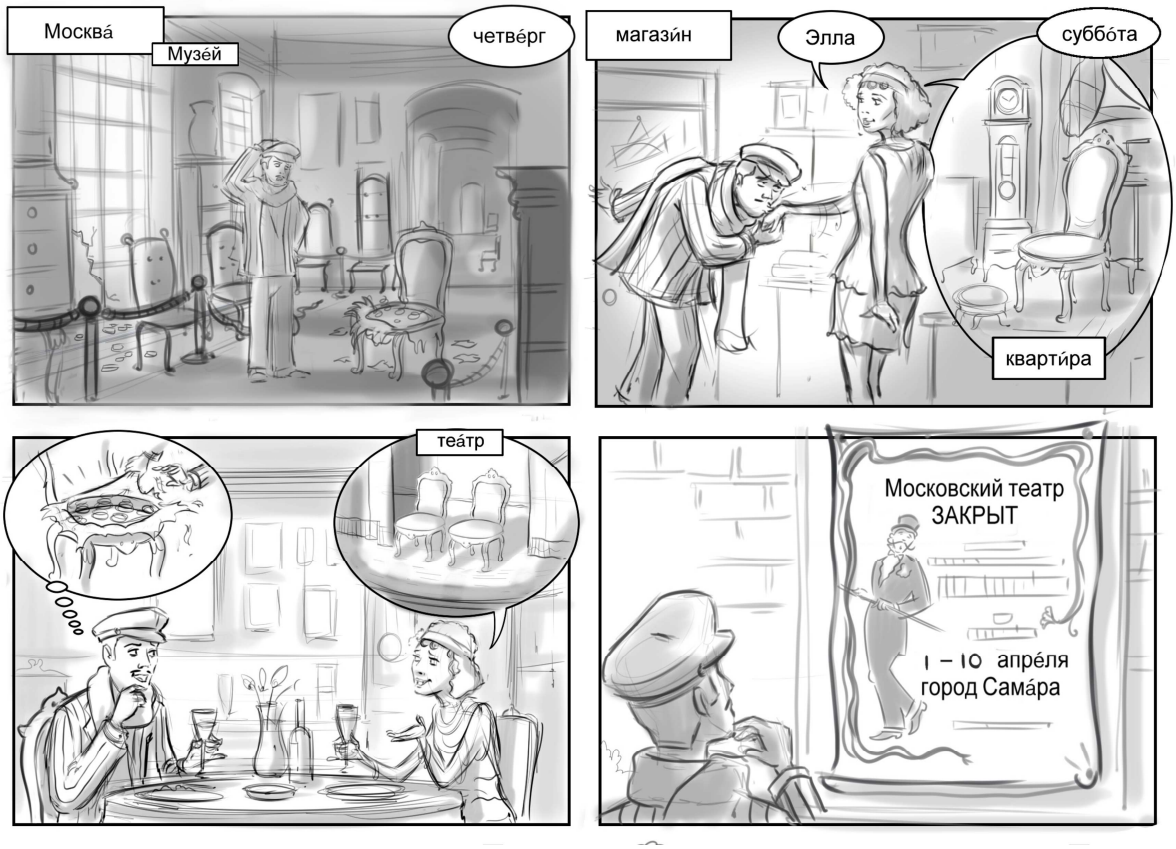
Following Cohn's (2014) suggestion that the future of the educational use of comics is in the "explicit manipulation of the component parts of the structure used in comics" (Cohn, 2014: 57), I have created my own comics for testing participants' oral production of case suffixes.

The comic strips for this Comics Test are based on a classic satirical novel "The Twelve Chairs" by two Soviet writers Ilf and Petrov (https://www.goodreads.com/book/show/158516.The_Twelve_Chairs ; https://www.e-reading.club/book.php?book=44438). It is set in Russia at the end of 1920s, where the main character goes on a hunt for treasure, hidden in one chair from an expensive set of twelve, sold one by one at auction. The plot, where he visits different places and meets different people, presents a context suitable for using both Prep. and Acc. case forms in the majority of the contexts listed in Table 11a. For the purpose of this testing, the plot has been heavily adapted. The 19 panels, produced by a professional artist from my sketches and image references, contain some captions with place names and time references (example in Figure 13), aimed to encourage learners to use the target forms.

Figure 13

Panels 9-12 of the Comics Test

(Captions in Panel 9 say “Moscow”, “museum” and “Thursday”; in Panel 10 – “shop”, “Ella”, “Saturday” and “flat”; in Panel 11 – “theatre”; in Panel 12 – “The Moscow theatre is closed. 1-10 April, the city of Samara”.)



The test panels were created so that they would provide a similar number of situations for both of the cases, and several panels in the Comics task were designed to make Acc. more or less unavoidable. Moreover, the Comics Task was designed in such a way that it would elicit as many instances of cases contexts from Table 13 as possible. However, in order to keep the amount of production manageable for the low proficiency level of the participants, the musical instruments and games contexts were not included in the Comics Task and were only produced in the Interview part.

Moreover, in order to address the issue of marking of Acc. masculine zero-suffix (which is syncretic to masculine base forms (see Table 2)) as an inflected case form (as opposed to the incorrect base form in Acc. case context), some of the Acc. case contexts were paired up with Prep. case contexts in the same lexical set. Thus, the opposition of case forms would help to see whether zero-inflection in Acc. is consciously produced to distinguish it from Prep. in a paired context. For example, games (Acc.) and musical instruments (Prep.) were used after the same verb *igrat’* (for “to play”), while both months (Prep.) and days of the week (Acc.) were time references. The most prolific pairing up was achieved with location (Prep.) and direction (Acc.), as these in addition to familiar vocabulary, included new place

names. For example, Comics Test panels were designed in such a way that participants would need to say *Stul - v Novgorod**e**.* (for "The chair is in Novgorod*[Prep. Acc. masc.].*"), as well as *On edet v Novogorod_.* (for "He is going to Novgorod*[Acc.masc.].*").

Furthermore, the new nouns were also used to investigate participants' ability to apply inflection to new lexis. These new place names were the names of Russian cities which were not used in the learning materials and participants were unlikely to have come across before. These were given as captions on comic panels and were only used location (Prep.) and direction (Acc.) contexts. This aimed to prevent the use of chunking in testing, when learners produce a pre-learned case form without realizing its constituents.

Moreover, every possible effort was made to elicit a reasonably balanced number of masculine and feminine case forms (see discussion of gender selection in 5.3.2.). However, it has been expected that participants would produce slightly more masculine forms than feminine, as month and the majority of frequently-used transport vocabulary in Russian, for example, are masculine (as are many of commonly-played games). In addition, the majority of names for places where students live (used in the Interview part) were also classed as masculine, e.g., *Lids* (for "Leeds") and *Jork* (for "York") (see more on gender distribution in 6.7.1).

The panel roughs for Comics were piloted with my RS group (during Year 1 of my PhD), revealing the need of two minor adjustments which made the plot easier to follow. As a result, two panels were edited.

Before the test, students were given a short introduction in English (3 sentences), describing the setting and the main characters, without uncovering the whole plot. Considering a very low proficiency level of the subjects, occasionally they were helped to navigate the plot by questions (in Russian), similar to those used in the interview test, for example, *Kuda on edet?* for "Where is he going?" or *Kogda on tam?* for "When is he there?". This kind of questions prompt a beginner with a verb or a context which would instigate the use of a case, without providing a required case form. (The number of questions asked within Part 2, varied, depending on participant's confidence in speaking Russian). This combination of picture prompts and interview techniques gave learners the opportunity to attempt sentences of their own, but at the same time eliminated a chance that a participant would not produce a sentence because they might have forgotten a word (with vocabulary not being tested) or cannot think of what to say. In addition, these questions sometimes helped to steer participants towards using Acc. where both, Acc. and Prep. structures were possible.

As the present study aims to investigate different factors influencing learners' processing of case inflection, following the discussion in 3.1, there was no time restriction and testees were allowed to go back to what they have just said and to change their mind (they were informed that it is the last option that would be counted). With fluency out of equation, these arrangements were to allow participants to self-correct and possibly illustrate their thought process, revealing what slows them down and requires more processing time. At the same time, these conditions are favourable for processing, making learners feel less pressure, thus enabling them to produce the best possible performance. As it happened, this has provided some valuable insights in how learners arrive to their decision on a particular case form (see 7.2. for discussion). Moreover, it has been demonstrated by research that the ability to modify their output leads to a more successful performance by learners (e.g., Gass, 2015: 185; also discussion in 3.3.3).

Furthermore, to ensure that the difficulties in vocabulary retrieval do not impede learners' performance on inflection, participants were to be prompted challenging vocabulary items. Similar prompts were used in DeKeyser's (1997) experiment, except DeKeyser presented all 32 vocabulary items on screen. During the piloting, it has become apparent that learners were struggling to remember two vocabulary items, essential for the comics plot – *dengi* for "money" and *teplohod* for "passenger boat". Initially, Test 2 was planned face-to-face and flash cards were made with most commonly required items, but when it was changed to the online format during the COVID pandemic, vocabulary prompts were presented via Skype chat.

To test experimental participants, it was decided to use both the interview test and the Comics Test, because the question-answer format of the interview was familiar to all participants and was expected to be comfortable, as question-answer activities were frequently included into language lessons; thus it would demonstrate what learners are capable of producing within their comfort zone. There was a danger, though, that students might produce pre-fabricated answers, as, considering a very low proficiency level of the participants, the range of questions was quite limited. In contrast, the comics test can elicit new phrases which could not have been rehearsed, due to the specific context; the unfamiliar format though, had the potential to put testees under pressure, thus affecting their performance. Also, in an interview test, low-level learners might replicate the examiner's language structures (priming effect (Anderson, 2015)), while the comics test will give learners flexibility in how to construct their sentences.

The Comics test was presented at the AATSEEL conference (American Association of Teachers of Slavic and east European Languages) and evoked a lot of interest. The test was initially shared on Open Science Framework website, recommended by home University, and now is licensed by the University of Leeds:

- https://licensing.leeds.ac.uk/product/speaking-test-1-testing-russian-grammar-in-beginner-learners-speech-comics-type-picture-panels
- https://licensing.leeds.ac.uk/product/speaking-test-2-testing-russian-grammar-in-beginner-learners-speech-comics-type-picture-panels

Each participant was recorded individually. The audio recordings were made on a Windows 10 laptop. Due to the COVID-19 restrictions, experimental participants were tested online, using Skype platform (see 5.3.1. for reasoning); though Zoom was kept as a back-up option and was used twice when there was a lot of background noise on Skype and when the Skype signal was weak. The comic strips were digitalized and presented to participants via Screen Share function. The tests were audio-recorded using the laptop, rather than online recording function, for reliability, which proved to be exceptionally useful, when, on one occasion, my Skype stopped working in the middle of the test, while the recording remained unaffected and was resumed after re-starting Skype; thus no data was lost. All 27 participants, who completed the instruction period, took Speaking Test 2.

The Interview part of Test 2, according to my research design, lasted five minutes. Then the recording was paused, while the Comics task was explained; this also worked as a welcome break for learners. The length of the Comics task varied (as the time was not restricted, see reasoning in 5.4.4.), with an average of 16.8 minutes and the two extremes of 12 minutes (P2008, one of the top WM spans) and 23 minutes (P2019, the lowest WM span).

As participants were not given any specific training in listening, though they listened to the teacher and to other participants in class, some testees experienced difficulties in understanding some questions (e.g., P2017, P2021), particularly at the start of the test. In order to enable those testees to have an equal chance of producing enough obligatory case contexts in their speech, some questions or particular words were typed for them in Skype Chat. These mainly involved *Gde vy živëte?* (for “Where do you live?”), *U nego/neë est ...* (for “S/he has …”), *segodnâ* (for “today”) and *obyčno* (for “usually”) - all four containing phonology perception challenges. No more than two full questions per test were typed up (more often no typing was required); the maximum number of individual words typed in Skype Chat because of listening comprehension difficulties within one test, was five.

Furthermore, as the testing did not aim to assess learners' vocabulary (which participants were informed about just before the test), if a testee had difficulty recalling a particular word, essential for eliciting a case form, the word was given to them in Skype Chat. However, to encourage participants to use their own resources to construct their sentences, instead of relying on vocabulary being given to them, they were explicitly discouraged from asking for words during the tests and were warned that if they do so, they might not be given a word, thus leaving vocabulary prompting at the discretion of the researcher. The vocabulary items were mainly given to elicit expected case forms. No more than five words were prompted in one test, with the most common words, typed in Skype Chat (*teplohod* for "a boat" and *poezd* for "a train"), both introduced in the last but one session of the Instruction period.

Considering low language proficiency level of the testees (see 5.2.1.), as well as the change to the online mode of testing during the COVID-19 pandemic (6.3.1.), there was no strict restriction on the use of English during the tests, though participants were discouraged to use English, if they could help it. On numerous occasions, participants commented that knowing that they could use English, if they needed to, made them feel more relaxed. The amount of English in the recordings varied a great deal and, I believe, depended on two factors: firstly, on technical issues (mainly on whether the testees could see pictures during screen share and whether the images were large enough for them to read the captions); and secondly, on how confident learners felt about their Russian. Some participants hardly ever adhered to English (e.g., P2008, P2020), while others felt the need to make comments in English and check that they understood my questions correctly (e.g., P2019, P2021). As it happened, some of those comments have proved to be very valuable, as they gave glimpses of how learners' thought process worked when they were choosing a case suffix. In addition, if I noticed that a participant was getting tired or tense, to give them a break and make them more at ease, I sometimes made a comment or a joke in English; they often laughed and got back on tracks. I had to be careful to keep the balance between distracting testees and giving them a break but I believe that I struck that balance right. The data show that my comments in English did not adversely affect testee's performance and in a few cases, learner's deteriorating performance improved after this kind of a break.

The difficulty, which could not have been foreseen, was that, during the COVID-19 lockdown, when participants could not go anywhere or plan for a holiday which was planned as one of the topics for the discussion in the Interview Part of Test 2 (which would have provided a range of contexts for Acc.), the choice of questions that could have been asked, for ethical reasons, was very limited. The verbs of motion for regular (repetitive) travel were not taught within the learning sessions, thus could not have been employed to adjust the

context. Consequently, the Comics Test, which did not depend on participants' personal circumstances, has proved to be of vital importance.

5.4.5. Delayed Speaking Test (Speaking Test 3)

Aiming to examine the long term effect of the instruction based on the proposed framework, another round of tests was carried out 6-8 weeks after the first post-test (Speaking Test 2), in order to compare whether participants' accuracy for the production of the Prep. and Acc. case suffixes, which were tested in Test 2, was retained.

In size and structure, the Delayed Test was very similar to Speaking Test 2, used at the end of the instruction period. It included two parts – an interview task and a comics task. Due to the limited language base, possessed by beginner learners, the questions for the interview part were very similar to those used for the interview task in Speaking Test 2, though the order and the exact phrasing varied slightly, depending on participant's answers. On the contrary, the comics task was a completely new test, to avoid task similarity priming effects. Though it was testing the production of the same cases (Prep. and Acc.) in the same contexts, it was using a new set of comic strips, based on a different story.

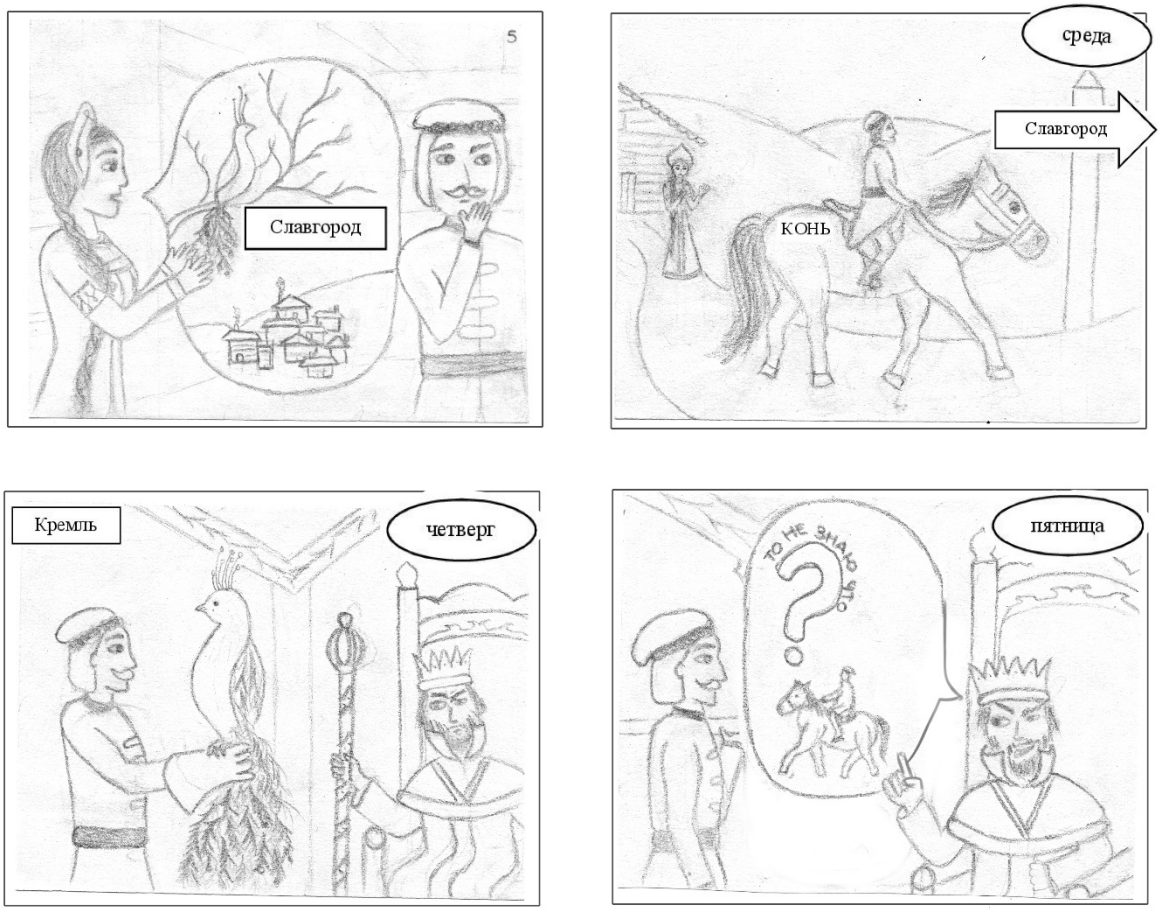
For the Delayed Test, I used a political satire play-poem "The Tale of Fedot the Strelets" by Soviet writer and actor Leonid Filatov, written in 1985, though the storyline of his poem is based on Russian folk tales, providing a lot of socio-cultural information. Due to the exhaustion of the funds available, the panels were drawn by two Year 11 students (free of charge), from my sketches and the image references that I provided. Their panels were then digitalised (see example in Figure 14) by a professional artist, to enable online testing due to the COVID-19 lockdown restrictions (see 5.3.1 and 5.4.4. for details). The difference in visual aesthetics of the two sets of comic strips was not significant and did not affect elicitations in any way.

The number of panels was exactly the same as in the first Comics Test used in Speaking Test 2 (19). The ratio between expected masculine and feminine case forms was similar (14:9 respectively). The number of expected case contexts for Prep. and Acc. were kept similar too (17:14 respectively). The RS group was used to pilot the Delayed Test.

The success of participants' case inflection production was measured by the number of correctly produced inflection in the obligatory case contexts (see more details in 6.1).

Figure 14

*Examples of **C**omic Panels for the Delayed Test (Speaking Test 3)*



In order to ensure the comparability of the data collected from the two tests, every effort has been made to ensure that the procedures of administering the test and recording participants' audio samples were the same as for Speaking Test 2 (see 5.4.4.). From 27 experimental participants who took the post-test, 14 took the Delayed Speaking Test, which is over 50%, required for valid comparison.

Considering the unpredictable nature of learners' production, I believe that the designed tests have provided a rich sample of required case forms.

5.4.4. Working Memory Test (WM test)

To explore possible reasons for learners' individual variation in ability to learn, process and produce case inflection, participants were asked to take a WM Test. Among various tests available, it was decided to test the WM operational span (Engle et al., 1992), as it was thought to be more relevant to the task of switching between different forms and functions of

inflection. WM's role in language processing - both as potential limit on intake and learning, also impacting on task difficulty in performance has been widely investigated (Conway et al., 2005 for a review).

Originally, the WM test was planned to be administered in class to each of the three experimental groups during Week 17. Due to COVID-19 restrictions, imposed during Week 17, it had to be completed online, independently by each participant. It took time to find a WM test suitable for these new conditions, consequently the test was done after the instruction period had finished.

Following the guidance of Conway et al (2005) on different characteristics of the WM span tests, the Operation Span Task (OSPAN) by Millisecond was chosen as the most suitable. It was accessed and conducted online, requiring a two-minute installation of an Iquisite Player, which did not cause any issues for accessibility for remote participation. The OSPAN test was about 20 minutes long, but also included a practice session, which was useful.

WM test was piloted on several volunteers, as well as on two types of devices, Windows 10 laptops, as well as two Apple laptops. It was discovered that it could not be stopped or paused, which, on one side, was very convenient, considering that participants were completing it independently, while, on the other side, it did not allow for any unpredictable situations. It was suggested to all participants that, should any disruption happen during the test, they should inform me immediately. Fortunately, all 27 participants completed the WM test without any interruptions, as firstly, none were reported and, secondly, the results were received within the time expected.

The detailed instructions were written and emailed to the participants in advance, so they could familiarize themselves with the procedures. The link to the Millisecond website though was only sent on the morning of the test, to exclude a possibility of participants practicing the test. Each of the experimental groups was also offered a Skype meeting at the end of the instruction period, where they could ask questions about testing procedures (including WM testing), their data and the results of the study, as well as getting advice about further learning, choosing a textbook and practicing their Russian. All three experimental groups opted to have a meeting, and 25 out of 27 learners attended.

The best effort was made to create similar conditions for taking the WM test for all of the participants. All 27 participants, who completed the instruction period, took it straight after their Speaking Test 2, which was scheduled within the same timeslot between 12.00 and

4.30 pm, but on different days. They were also encouraged to have had a good sleep the night before and to eat a meal at some point before testing, to avoid variations due to hunger and tiredness which are known to impact WM performance. The results of the tests were emailed to me by each participant straight after they completed the test, that is, within 30 minutes of the completion of Speaking Test 2.

During the OSPAN test session, participants were presented with letter sequences, ranging from 3 to 7 letters, and were asked to recall the letters, as well as their sequence after each presentation, by selecting letters from a letter matrix. During the recall stage, each letter in the sequence was preceded by a maths problem (e.g., (8*2) - 8 = ?) followed by a proposed solution ("e.g., 9") - participants were asked to judge whether the solution is correct or not. A test session included 15 trials – that is, 3 repetitions of 5 set sizes, whose order was random. The outcome would yield a WM Absolute Score, as well as Total Correct score and Maths Accuracy Errors. The results were recorded in a table.

5.5. Qualitative investigations

The supplementary qualitative component of the current study aims to answer the same overarching question of whether spiralling is an effective approach to teaching Russian case inflection (see details in 5.1.1); my qualitative investigations though, explore it from a different angle. Their purpose is twofold – first, to examine how learners perceived the instruction which was based on the proposed spiralling framework; and second, to draw out learners' perceptions of and possible reasoning behind learners' difficulties or successes in mastering Russian case inflection (that is using it in speech). More specifically, I look at which aspects of language learners focus on, whether they notice case inflection, how much attention they pay to grammatical form, as compared to lexical meaning, as well as possibly the role of processing load in the effectiveness of learners' case production (see discussions of each aspect in Chapters 2 and 3). The idea is that the analysis of these aspects would contribute to identifying more challenging areas and possible ways of facilitating inflection production process, as well as seeing whether the proposed framework addresses these.

Furthermore, I have added learning strategies to my investigations, as these have been the focus of research which looks for means of facilitating learning, and this is a long term goal of this project. Oxford (1990) defines learning strategies as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferrable to new situations" (Oxford, 1990: 8). Chamot (2005) points two main reasons for researching learning strategies: firstly, they provide insights into learners'

cognitive processes, and, secondly, other learners could benefit from acquiring newly-discovered strategies; both aspects being the object of my investigations. Fully aware that learning strategies constitute the research area in its own right, and considering that the present study focuses on cognitive aspects of language learning, I have decided to examine only cognitive learning strategies – that is those which involve "manipulation and transformation of the target language by a learner" (Oxford, 1990: 43).

The qualitative component of this study includes learning diaries, questionnaires and interviews. All three instruments are based on the principle of self-reporting, which is considered to be the only method which can help researchers understand the way learners process linguistic information and use it to produce their own language (Chamot, 2005). However, self-reporting has some considerable limitations, as it depends on learners' ability to accurately analyse their thought process, give an account of their decision-making stages and pinpoint the causes of their actions. Moreover, beginner learners do not necessarily possess sufficient linguistic knowledge to be able to identify the language aspects which would be important for the study. As frequently quoted Grenfell and Harris summarised with regard to investigating mental processing: "We work with what we can get, which, despite the limitations, provides food for thought" (Grenfell and Harris, 1999: 54).

That is why, I decided to combine the three tools, so that they could complement each other and enable me to create a more accurate picture of how learners attempt to map the explicit information that they receive, onto the case forms that they produce. As a result, unstructured learner diaries have provided data, which was mainly not influenced by researcher interference, but the amount and the quality of data were subject to being compromised, while questionnaires direct learners specifically towards the aspects of case inflection production, examined by this study, but are fairly structured, thus restricting learners' choices. Finally, interviews allowed me to build on all previously collected data, from both the quantitative component and from the previously-used qualitative tools, exploring the same aspects at a deeper level, thus providing some kind of triangulation within the qualitative component. Each of the tools is described in detail below.

5.5.1. Learner Diaries

The first qualitative instrument used in this study is learner diaries (or learner journals, or even learner memoirs) as a type of introspective personal narrative (e.g., Ellis, 1985), which is considered one of the most effective tools for investigating "what goes on in learners' minds" (Jones, 1994: 443) and a way of collecting data about learners' mental processes

(Ellis, 1985). A diary provides the learner's account of learning events, as well as their perceptions of these events and reflections upon them. Oxford (2011) differentiates two components of a learner diary, namely, external (factual, observable) and internal (interpretive, reflective), stressing the subjective nature of their content. I believe that here subjectivity might be an advantage, providing researchers with a learner's perspective of the learning process, as Janesick (1999) noted, "giving participants an active voice" (Janesick, 1999: 522). Among the main themes of learner diaries, Pavlenko (2002) lists learner's difficulties, their achievements and their failures in language learning, which is exactly what the qualitative investigations of this study aim to elicit. Ma &Oxford (2014) also add learning strategies to the list, which are also examined here. Faerch & Kasper (1987) point out that learner diaries are beneficial for analysing the status of learner's declarative knowledge, which, in the context of the current study, appears useful for drawing comparisons between participant's actual performance in speaking tests and their diary entries.

Another important characteristic of learner diaries, reported in literature, is that learners can only record the language aspects which they are aware of (Ellis, 1985) and, considering that recruited participants did not have any reasonable background in languages, there was always a risk that they might not necessarily focus on the complicated phenomenon of case inflection. Thus, initially, I had a choice of whether to direct learners towards monitoring case suffixes, by making the diaries more structured and providing more detailed guidance, or by leaving the narratives unstructured, in order to see whether participants would concentrate on case inflection, and, if yes, how they would go about it, but then running the risk of data being scattered. As the former would have been fraught with danger of making participants aware of the phenomenon studied, thus causing them to try harder with case suffixes than they would have done otherwise, or apply strategies that they would not have done independently, and consequently ruining the authenticity of the account, I decided to leave learning diaries unstructured, asking learners to reflect on what was difficult, why they thought it was and what they did to overcome problems, as well as what made their learning easier and why. In addition, participants were encouraged to put down their suggestions of what could have helped them learn a particular aspect (though the aspect was not named).

During week 4 of the Instruction period, when most of the participants were fairly settled into their learning, they were sent an email calling for volunteers to keep a journal. As a result, four participants responded. They were given a blank copy-book to use (with a Russian related cover), but were also offered an option of keeping their diary in an electronic form (e.g., on a computer). All four volunteers opted to write their notes in the book.

Volunteers were asked to keep a regular record, at least once a week, for example, after each session. Considering that the participants were recruited from non-language departments and were not receiving any credits for this course, there was no guarantee how much time they would devote to this activity, thus how detailed their entries would be and whether they would even complete their diaries (the problem was also reported by Janesick, 1999).

5.5.2. Qualitative Questionnaire

The second instrument, which I have selected for my qualitative component, is a questionnaire. Due to the specificity of the subject of my investigations, namely, Russian case inflection production, I could not identify any questionnaires, devoted to this topic, in the research literature. Thus, I have designed a new tool, informed by standard practices as best identified through the Strategy Inventory of Language Learning (SILL) (Oxford, 1990), as detailed below.

My Qualitative Questionnaire was semi-structured, having a mix of closed and open questions, which require information related to case inflection, while at the same time giving participants an opportunity to expand on their initial answers or comment with no answers suggested, if they wished. This Questionnaire was split into two parts, which investigated case inflection from different perspectives and at different levels.

Part 1 (Appendix G) examined learners' perceptions of difficulties in learning different aspects of Russian (including case inflection) within the instruction period, enabling me to see where learners place case suffixes in comparison to other language aspects studied, on the scale of degrees of difficulty. This part also enquired about how participants perceived the spiralling instruction that they had received. After switching to the online teaching in Week 17, the section was added, which examined how participants perceived the effects of the change of the instruction mode (Question 5 in Appendix G). Part 1 was anonymous and was completed online on the Jisc website (https://www.jisc.ac.uk/), recommended by the home University. The format of the questions varied to suit the question content, as well as possible answers – multiple answer, multiple choice, scale and free text.

Part 2 (see Appendix H) was also semi-structured, but was dedicated solely to psycholinguistic processes of morphology acquisition, nominal case inflection in particular. It examined noticing, understanding of forms and functions, explicit/implicit aspects of learning, difficulties of processing and producing case suffixes, as well as strategies of choosing a

case suffix and cognitive learning strategies. In designing this second part of my Questionnaire, I used SILL, as a model for some questions (Oxford, 1990), which is extensively employed to collect data in learning strategy research and is considered to be a standardized measure. For example, I borrowed its well-attested grading scale for Questions 4, 6 and 9:

1 - never or almost never true of me;
2 - generally not true of me;
3 - somewhat true of me;
4 - generally true of me;
5 - always or almost always true of me.

Also, similarly to SILL, the choices were formulated using phrases in 1st person Singular, e.g., "I tend to analyse why the suffixes are there". Finally, a mix of answer choices from different themes of this investigation (processing load, explicit-implicit relationship, form-meaning relationship, cognitive strategies), were included within *one* question, with the aim of not making it obvious for learners what answer was expected. Table 14 presents an example of the split of the choices, which are introduced in a random order within Question 9 (see the actual ordering in Appendix H). (The other questions included multiple choice, one Yes-No item and one ordering entry.)

Table 14

The Split of the Choices, which are Introduced in a Random Order in Question 9.

9) If you believe that you do not always use noun suffixes when you SPEAK (or do not always use them correctly), what do you think the reason/s is/are?

Processing load	***Explicit vs implicit***
By the time I remember the words that I need, it's too late for me to think about suffixes.	I don't think I quite understand which suffix should be used when and say them randomly.
I miss more suffixes when I have to remember what I am supposed to say as well (e.g., reporting back on pair work).	I can explain which suffix I need in which context, but when I speak, they get confused in my head.
I think I am OK when there is one suffix in a sentence, but when there are more, I get lost.	I feel I get the first suffix ("Е" in "в Лондоне") correct more often, but haven't got used to using the second set ("У" in "в Москву") yet.

We now have so many different contexts when an suffix can be used, that I can't choose between them.	If I make a mistake, I often know what I should have said.

Cognitive strategies	***Form vs meaning***
I tend to say the first suffix which comes into my head.	I believe that, if I don't use the suffixes right, Russians would understand me anyway.
I get suffixes correctly in the phrases that I have learned.	When I speak, I have an English sentence in my head and make sure that I have all the corresponding words in my Russian sentence.
I believe that verb might help me decide which suffix the noun needs, so it's easier with a verb	I think that I don't have to have all suffixes perfect to be able to communicate, but I do need to get at least some of them right.
Before I say something, I try to think ahead a bit to work out which suffixes I need, and occasionally get it wrong.	I think that Russians would struggle to understand me if I don't use the suffixes correctly.

Part 2 was *not* anonymous, as the intention was to match participants' answers with their performance in Production tests, as well as to use these data for developing the individual interviews. Both parts of the questionnaire were distributed at the same time, after the instruction period had finished.

The best effort was made to keep a reasonable balance between collecting enough relevant data and not requiring participants to spend an excessively long time to fill the questionnaire in (see Dörnyei & Taguchi, 2009; Adams, Fujii & Mackey; 2005; on risks and limitations). For example, in Part 1 (administered through Jisc website), the closed questions were made compulsory for proceeding through the Questionnaire, while all free text comments and some open-ended questions were optional, in order to give participants a choice of how much time they would want to spend on this questionnaire.

Part 1 comprised 10 questions, which fit into 2.5 pages of A4 in Word, though on the website, they appeared on 5 pages, due to the format of the sections split. For example, there were separate pages for the Privacy Note and the concluding Thanks; the remaining

three sections were named “Before you started learning Russian”, “During your learning sessions” and “After you completed your learning sessions”, which I believe were more participant-friendly than a continuous list. Part 2 included 12 questions, which occupied 4.5 pages in Word, due to the amount of detail. The questions were divided into those related to explanations, reading and speaking.

For this questionnaire to be piloted, I needed volunteers who would have been learning Russian for at least 20 weeks. The RS baseline groups participants suited this criteria and both parts were sent out to them a week before being distributed to experimental participants. There was a fairly good response rate, as 6 out of 8 RS participants completed the draft version of the questionnaire. At the very start, it was discovered that one of the questions in the online part (scale Question 6) did not work as intended, as it was not set up correctly. That was corrected immediately and functioned properly for the rest of the piloting and when the questionnaire was administered to experimental participants.

5.5.3. In-depth Interviews

To “enhance the interpretability” (Mark & Shotland, 1987) of the data collected by questionnaires and through learner diaries, as well as to get deeper insights into the reasons behind some of the answers, I conducted six in-depth interviews. For obvious reasons, I planned to interview those who kept the learner diaries, as they provided more data, which were worth exploring further. Furthermore, as I am looking at individual variation, I selected two more learners, to be interviewed, from the opposite ends of the achievement scale, hoping to identify some noticeable differences in their approaches to inflection production challenges. Altogether, I recorded six in-depth interviews at the end of the teaching cycle.

The questions for the Interviews were developed from the relevant entries in Learner Diaries and from answers and comments in participants’ Questionnaires (see Green et al., 1989, on purpose of development). Thus, the In-depth Interviews were further exploring the aspects of explicit information processing, noticing of case suffixes, decision-making during inflection production and cognitive strategies in mastering case inflection, investigated by Part 2 of my Questionnaire. Unlike the questions in the Questionnaire, interview questions required extended responses – for that reason all questions were open-ended.

The basic questions were similar in all interviews, while some details and some follow-up questions varied, reflecting specific individual responses or comments, previously provided by the interviewees (as recommended by Gillham, 2000). Consequently, six interview

protocols tailored for each interviewee were compiled (Johnson, 2014) (example in Appendix I). One version of the interview was piloted with a participant from an RS baseline group.

Though I compiled a list of questions to ensure feasible thematic coding, I kept the interview as flexible as possible, in order to be able to respond to possible emerging directions offered by the interviewees. To help participants clarify or allow them to elaborate on their initial answers, probes and prompts were often employed. Johnson (2014) classifies this type of interview as an interview guide approach. All In-depth Interviews were recorded after the end of the Instruction period.

The format of qualitative data collection was partially adjusted, due to the restriction of COVID lockdown. For example, Learners' Diaries could not be handed in to me as a book and were scanned and emailed to me as PDF documents, which were later typed up into Word documents for the ease of coding (the original versions were retained for future reference, in case a need arises). Part 2 of the Qualitative Questionnaire was emailed to participants as a Word document after the learning sessions were finished. They were asked to fill it in and return it as an email attachment. Part 1 of the Qualitative Questionnaire had been planned to be conducted through the Jisc website. However, an additional question exploring the effects of the transition to the online teaching due to COVID lockdown, was introduced (Question 5). Finally, the In-depth Interviews, which were initially planned to take place face-to-face, were conducted online, using Skype, which was used during the online learning sessions and was the platform familiar to participants (see 5.3.1 for reasoning).

Overall, the collection of quantitative data for this study mainly went as it was planned. The only change which was unavoidable was switching to the online mode of testing for Test 2 and Test 3, as well as for WM test. Despite some of the challenges, 27 from the 28 experimental participants who started the second (case-learning) half of the teaching intervention have completed Test 2 and the WM test, with over a half of them taking Test 3. No technical issues resulted in any loss of data, which was successfully transcribed, marked and summarized, ready for the statistical analysis.

5.5. Conclusion

In this Chapter I have put forward my Research Questions and outlined the design of the study which aims to answer these questions. By combining quantitative and qualitative research methods, I produced a sequential mixed method study involving recruiting volunteers for my longitudinal teaching intervention and collecting case inflection production

data in three rounds of tests. In addition, I have collected qualitative data, using learning journals, questionnaires and selective interviews. This has allowed me to analyse the phenomenon of case inflection production from various perspectives, including learner variation. The overall scheme of data collection is presented in Table 15.

Table 15

Data Collection Times, Procedures and Instruments.

	Data collection instrument	What is tested	Which RQ/s answered	When administered	How many participants
QUANTITATIVE	Test 1 (interim test) "Guessing Game"	Accuracy of production of Prep. "initial" suffixes for masculine and feminine in one case context (location); whether it is different for different genders;	RQ2 RQ5 (regarding gender)	Week 15 of instruction period	28
	Test 2 (post-test) Part 1 – Interview; Part 2 – Comics Task	Accuracy of production of Prep. and Acc. "initial" suffixes for masculine and feminine in different case contexts (see Table 13); to investigate the accuracy drop for Prep.; the differences in participants' performance for the two cases; the effect of the three factors in RQ5; the correlation with WM	RQ1, additional data for RQ2; RQ3, RQ4, RQ5, RQ6,	immediately after the instruction period was completed	27
	Test 3 (delayed test) Part 1 – Interview; Part 2 – Comics Task	Retainment of accuracy of production of Prep. and Acc. "initial" suffixes for masculine and feminine in different case contexts (see Table 13)	contributes to RQ1, RQ2, RQ3, RQ4, RQ6; also RQ5 (regarding gender)	6-8 weeks after Test 2	14
QUALITATIVE	WM span Test – Operation Span Task (OSPAN) by Millisecond	WM operational span	RQ6	immediately after Test 2 was completed	27
	Learner Diaries	Retrospective unstructured, participants' account of their learning	RQ7	filled in from Week 4, collected after instruction period	4
	Qualitative Questionnaire Part 1 (anonymous),	Evaluation of spiralling instruction, perception of difficulties of learning Russian, effects of COVID	RQ7, contributes to RQ1, RQ2 and RQ3	after the instruction period was completed	24

	semi-structured				
	Qualitative Questionnaire Part 1 (not anonymous) semi-structured	Participants' perception of learning Russian case inflection specifically	RQ6, RQ7	after the instruction period was completed	21
	In-depth Interviews, semi-structured	More elaborate narrative on interviewees' responses in Qualitative Questionnaire and Learner Diaries (if applicable)	RQ6, RQ7	after Test 3	6

Chapter 6
RESULTS OF THE STUDY

In this chapter, I will present the results of my empirical investigations, conducted, in order to answer my RQs, posited in Chapter 5 (see 5.1.1).

6.1. Preparation of quantitative data for statistical analysis

Three speaking tests (see Table 15), administered to the teaching intervention participants, provided me with the data that were used to answer the majority of my Research Questions (see 5.1.1), namely, RQ1, RQ2, RQ3, RQ4, as well as RQ5 and RQ6. In this section, I will outline how the recorded oral production samples were transcribed, marked and summarized, to enable the creation of three statistical data sets.

6.1.1. Transcribing oral samples

For the purpose of the statistical analysis, the raw data from the audio test recordings were transcribed into Word documents. The transcriptions were written in Cyrillic, as that was considerably less time-consuming for me, as I am a native Russian speaker. The excerpts selected for illustrative purposes, were later transliterated into English with the help of "Tranliteration" online software programme, which can be found here: https://www.translitteration.com/transliteration/en/russian/iso-9/.

However, finding appropriate Voice-to-text (or Speech-to-text) software proved to be very challenging, as the software available, if set for Russian, appeared to recognise only native speaker's Russian and not identify any of learners' speech. The alternative of using English as a transcription language was also unsuccessful, as, due to different phonological structure of English, Russian words produced by learners, were split inappropriately in the outcome, while different vowel quality programmed into the software rendered them unrecognisable. After further online research, I selected Live Transcribe, a highly sensitive application, designed for deaf users, released by Google in February 2019. It is listed on several websites for deaf or hard-of-hearing people, e.g., https://www.hearinglink.org/living/loops-equipment/useful-apps-for-hearing-loss/
In addition to picking up some percentage of learners' responses, it has provided a dialogue layout (rather than a solid text, commonly produced by other software), which has proved to be a more efficient output format to work with. For example, in the Test 2 and Test 3 interviews, participants' answers, which were not registered, could be filled in between

interviewer's questions, which considerably sped up the transcription process. In the future, I would use Live Transcribe software while recording the tests, which would save considerable amounts of time by providing a rough text version of a recorded audio straight at the end of the recording. Despite the above measures, over a half of learners' speech was transcribed manually. As participants' pronunciation varied, the software registered more speech for some learners than for others. On average, I was managing to accurately transcribe two samples of Test 2 or Test 3 a day; each test of between 4000 and 5000 characters (with spaces).

Test 1 samples were considerably shorter, but as I was not involved in the "Guessing Game", the amount of speech, picked up by Live Transcribe was significantly smaller. Another reason could possibly be that learners' pronunciation was less distinctive in Week 15 than in Week 21, thus rendering less eligible lexis. Therefore, the use of Live Transcribe for Test 1 was abandoned, and all samples for the first oral test were transcribed manually.

Each sample was saved under Participant's number, thus no personal information was included into any of the transcriptions.

6.1.2. Marking of transcribed oral samples

After all the transcriptions were ready, relevant case inflections were marked, initially by myself and then selectively by the second rater, who was a native Russian speaker studying at home University (see 5.2.3 for details). The aim of this procedure was to identify all relevant inflection contexts for the two cases investigated (Prep. and Acc.), as well as rating the identified instances as correct or incorrect (further differentiations of error types were made at the summarizing stage).

For this purpose, a rather straightforward marking system was adopted: correctly assigned suffixes (including zero suffix) were marked in green; incorrect suffixes (of any type) were marked in red. That allowed the lay second rater to efficiently mark 3 samples randomly selected from each of the three tests; 9 samples from the total of 57; that constituting 15% of data. To determine inter-rater reliability of the marking, Cohen's Kappa coefficient was calculated. The Kappa value was 0.89, which is interpreted as "strong agreement" and is very high. A rather limited number of disagreements involved the Prep. form of *muzej* (for "museum"), namely, *v muzeje* for "in the museum *[masc Prep]*" in Test 2. The combination of the three sounds at the end of the form ("eje") is challenging for English speakers to pronounce, as it does not comply with the rules of English phonotactics, that is, how sounds

combine in English. Thus it was often difficult to differentiate between the Acc. zero-suffix form ("v muzej") and the Prep. inflected form (*v muzée*, pronounced as /m u z é j e/), with the last vowel being reduced, indistinctive or hardly heard. After discussing this issue with the second rater, it was decided that the word *muzej* would not be included into obligatory case context. This has not inversely affected the balance between the two cases, as *muzej* was predominantly used in Prep. (with no instances of Acc. contexts), and initially there have been more predicted contexts for Prep., than for Acc., with more Prep. case forms being actually produced (see 6.3.1). Also, as *muzej* is a masculine noun, and more masculine case forms were expected to be produced in the tests than feminines, the gender balance was not affected either. Thus, the absolute majority of my marking was deemed reliable and the remainder of the tests recordings were marked by myself.

To make sure that the marking criteria is consistent throughout, a marking protocol was developed. Participants' responses were coded as correct if they produced the correct case form for the gender and one of the case contexts listed in Table 16 (see reasoning for context selection in 5.3.2). Only these case contexts, which were practiced during the Instruction period and were aimed to be tested, were marked, rather than any case context.

Table 16

List of Case Inflections and Case Contexts for Marking

Prep.	Inflection (for both genders	Acc.	Inflection	
			m (inanimate)	**f**
Location	"e"	games	zero	"u"
musical instruments	"e"	days of the week	zero	"u"
Months	"e"	direction	zero	"u"
Transport	"e"	direct object	zero	"u"

The choice of preposition was not taken into consideration, because the two prepositions which are used for the two cases (*v* for "in"/"at"/"to" and *na* for "on"/"at"/"onto"), do not determine the choice of inflection (see more in 4.1). Thus, the case suffix appropriate for the context, was marked as correct, even if the wrong preposition was used before the noun.

In Test 1, a very limited number of Nominative forms was examined briefly during qualitative analysis of the samples for instances when participants were coming to grips with Function assignment. Only incorrect instances of Nominative were marked, as the absolute majority of nouns in Nominative contexts were inflected correctly. The use of Nominative in "I have …"

constructions (which are reversed in Russian, e.g., “A pen is owned by me”) and Present predicate (e.g., “ I am a student”) was not marked.

The biggest issue was marking of Acc. masculine zero-inflection, which matches the masculine base form in Russian. Addressing this was built into the Comics test design by pairing up of case contexts from Prep. and Acc. (see 5.4.4). Thus, the zero-inflection in Acc. contexts (directionality, games and days of the week) was only marked as “correct” if it was produced in opposition to overt Prep. masculine suffix (in location, musical instruments or month contexts). In some individual cases, when, for example, months were not acquired as a set (e.g., P2023), the production of Acc. masculine in days-of-the-week context was cross-checked against feminine days of the week (that is, Acc. masculine was only counted as correct, when correct Acc. feminine was produced).

As testees were allowed to make several attempts at getting a case suffix (as discussed in 5.4.4), though only the last attempt was counted, all other attempts were marked in brown and recorded as “attempts”. A case form was considered an “attempt” if it was produced by a testee as part of a close sequence of forms of the same noun for the same context within the same sentence. For example, *Kuda? -V Samar**e**[fem Prep].... Edet v Samar**a**[fem base form] .... v Samar**u**[fem Acc]* “ (“Where (to)?” – in Samara*[fem Prep]* …. Going to Samara*[fem base form]* …. To Samara*[fem Acc]*”) (P2025). I believe that these forms, produced by a learner before the final decision is made, offer some insights into learners’ thought process in choosing a case suffix (see 5.4.4 for discussion). At the summarizing stage, “attempts” were grouped into several categories (see 7.1.4).

Furthermore, from very early in the process, it has become apparent that, in Test 2 and Test 3, there are quite a number of repeated case forms (both correct and incorrect). As in the present study, data is analysed qualitatively, as well as quantitatively, these repeats were considered an important characteristic of learners’ production process. On one hand, some learners had a persistent tendency to repeat what they have just said or heard, possibly to assist processing (e.g., P2001, P2003). Also, some other repeats were generated by the suggested context, for example, by questions asked in the interview, e.g., *Gde vy živëte?* (for “Where do you live?”) and *Gde vy rabotaete?* for (“Where do you work?”), which often, but not always, yielded the same answer *v Lids**e*** (“in Leeds *[fem Prep]*”)), or by the story line, used for the pictures (for example, a re-occuring event of discovering that there is no money hidden in the chair, eliciting Prep. form *v/na stul**e*** (“in/on the chair”)). On the other hand, it has been noticed that not all case forms are produced correctly even in an absolutely identical context, if the instances of that context are separated by other case forms, for

example, Participant P2006 produced correct Acc. (*idët v teatr_* (“is going to the theatre *[masc Acc]*”)) for Panel 11, but incorrectly used Prep.(*idët v teatr**e**** for “is going in the theatre*[masc Prep]**”) in exactly the same context for Panel 13. Similarly, Participant P2024 used *v avgust**e*** (“in August *[masc Prep]*”) for Panel 14, but failed to inflect the same month for Panel 18. These examples indicate that the processing conditions can be different, even if the linguistic context is identical. Thus, it is possible to speculate that at least some of these repeated forms might involve the same or similar mechanisms as constructing a new form which was not produced before (Function Assignment and Constituent Assembly), and might be worth analysing separately. For this reason, all instances of the same case forms were initially recorded as “repeats” and marked in purple (rather than in red or green). If a “repeat” was preceded by an “attempt”, it was always recorded as a separate entry and marked in red or green.

There were some smaller-scale decisions which needed to be made with regards to marking. Firstly, how to mark a combination of a proper place name and a generic word, such as “a city/town” or “a village”, e.g., *gorod Samara* (for “a city of Samara” in the Comics task of Test 2). In Russian, there are two options which are acceptable – inflecting both words or inflecting the generic word. As this type of structures were not practiced during instruction period, learners produced various combinations of forms. Having in mind that usage was not assessed, I decided to count one case form, whichever correct, and even if two are produced. For example, in the following three Prep. instances: *v gorod**e** Samara* (P2009), *v gorod**e** Samar**e*** (P2012) and *v gorod Samar**e*** (P2010, P2018, P2020), one correct inflection was recorded (even though the latter would not be acceptable in Russian from the usage point of view). However, a more challenging decision was generated in P2010 sample, in the following exchange: *Kuda edet teatr? - V gorod Samar**e***. (“Where is the theatre going? – To the city *[masc Acc.= base form]* Samara *[fem Prep]*”). The first word of the apposition (*gorod*) was produced in the base form, which matched the required Acc., but as the proper name (*Samara*) was put into Prep., the whole phrase was counted as incorrect, assuming that the first word was not intended as an Acc, but remained as an uninflected base form. Following the same principle, in *vidit podrug**u** Ell**a*** (“sees a friend *[fem Acc]* Ella *[fem base form]*”), one correct Acc. was recorded, despite the combination of the two forms not following the rules of Russian usage.

The same marking was applied to the scripts of all three tests, enabling me to summarize all relevant inflections for each of the test in one document, ready for statistical analysis.

6.1.3. Summarizing of quantitative data

The summarizing of the marked scripts of the tests recordings was carried out in three stages: I) extracting and tagging relevant inflection forms; II) counting inflections for each tag producing raw scores summaries; III) compiling overall summaries of success rates for each category. The data from the three tests were recorded separately, producing three separate data sets.

To extract relevant inflection forms, all marked instances of inflection for each test were copied from the individual scripts in Word and pasted into an Excel spreadsheet for that test. To provide sufficient information for qualitative analysis, case forms were copied within the relevant linguistic context (see Appendix M for an example).

After that, each case form was tagged, using Data Validation function in Excel, as this prevented any typos, which could complicate data analysis with the help of a software. Each instance of relevant inflection was tagged for:
Tag 1 - gender of a noun in an obligatory case context;
Tag 2 – case for the obligatory context;
Tag 3 - case context, in which it was produced;
Tag- 4 - whether correct for the context;
Tag 5 - type of error, if incorrect;
Tag 6 - attempt/s, if produced;
Tag 7 - test part where the inflection was produced (for Test and Test 3);
Tag 8 – degree of familiarity of the noun;

Below, I will provide some details for each of the tags listed above. The example of an Excel summary can be found in Appendix M.

With regard to gender (Tag 1), each case form was marked for one of two genders, namely, masculine (m) and feminine (f) (see reasoning for excluding neuter in 5.4.1). I have decided not to use the term “declention type” or “noun class” (see 4.1), which are often used in Russian case research (e.g., Cherepovskaia et al., 2021, Rubenstein, 1995) and can include more than one gender, because part of this study is related to the acquisition of case inflection, thus the term “gender”, acknowledged in SLA, appears to be more appropriate. Also, only one gender for each declension type was tested. All case forms in this study are Singular, as Plural category and Plural case forms were not taught within the Instruction period. Thus, the Plural noun *den'gi* (for “money”), essential for Test 2 storyline and

occasionally figuring in Acc. object context, was not counted. Besides, it has a Plural form synchretic to Nominative, thus not indicating the change of case.

Tag 2 involved mainly cases: Prep. (Prep.) and Acc. (Acc.) for Test 2 and Test 3, as the aim of these two tests was to test the accuracy for these two cases. For Test 1, only Nominative and Prep. cases were recorded, as, according to the research design, in Test 1, Acc. was not tested (see 5.4.1 and 5.4.3). No other cases were taught during the Instruction period or produced in the data.

The full list of case contexts for Tag 3, is provided in Table 13. There were two lexical items (used in Prep.) whose context was not straightforward, that is *na karantine* ("in/during the lockdown") and *na komp'utere* ("on the computer"). These were not planned to be introduced initially, but emerged as necessary for learners to make relevant sentences during speaking activities in the last four sessions since the start of the lockdown (see 5.3.1.) and were actively produced in the Interview part of Test 2 and Test 3. As *na karantine* was used in Prep. and as a time reference, it was tagged as "month", while *na komp'utere* was classed as "musical instrument", as the closest possible context in the corresponding case.

For Tag 4, the definition of "correct" is similar to that in 6.1.2. - that is, the nominal case inflection was considered to be correct if it is apprpopriate for both the gender of the noun-carrier and the case context, in which the noun is used, and correctly attached to the noun's stem, creating a fully-fledged case form (e.g., the feminine base form suffix "a" is replaced). All other inflections were recorded as "incorrect". The forms which were recorded as "incorrect", were also tagged for error type.

Initially, three types of errors were marked by Tag 5:

- base form, inappropriate for the context, that is, for Prep. or for Acc. fem, that require inflection ("base form"); ("Primary Form" in Rubenstein, 1995)
- case form, inappropriate for the context, rendered two types:
  if Prep. suffix "e" was produced in Acc. context ("subst by Prep.");
  if Acc. fem suffix "u" was produced in Prep. context ("subst by Acc.");

In addition, during summarizing one more type of error has emerged, when the inflection was not fully-attached to the noun stem, that is, there is a time gap between the stem and the case inflection or if the feminine base form suffix "a" was not taken off. After discussing this issue with the supervisors, it was aggreed to record this as a "postponed inflection error". However, as the number of errors was quite small, they were only used for qualitative analysis.

Tag 6 was originally created to record attempts, which are types of output modifications (see 3.3). For Test 1, two types of attempts were recorded, which are similar to the types of errors produced in Test 1, namely, base form attempts and postponed inflection attmepts. For Test 2 and Test 3, the number of attemps was insignificant and was not used in the analysis. Similarly to error types, these were used in qualitative analysis.

Tag 7 was only used for Test 2 and test 3 and marked a part of the test where inflection was produced, that is Part 1 ("Interview") and Part 2 ("Comics Task").

As familiarity of lexis was identified as one of the factors to be investigated (see 5.2.1), Tag 8 was introduced to mark each inflected noun as

- familiar ("familiar"), that is, introduced and frequently practiced during the instruction period, e.g., *universitet* for "University", *Moskva* for "Moscow";
- rare ("rare"), that is, introduced during the instruction period, but practiced for very limited time, e.g., *kassa* for "box-office"; or introduced very late in the instruction period, e.g., *poezd* for "train";
- new ("new"), that is, not introduced during the instruction period and not encountered during the practice, e.g., *Novgorod* (a name of a Russian city), *Samara* (a name of a Russian city). These items were only present in the Comics Task of Test 2 and Test 3, and were given as captions.

Tag 8 was not used in Test 1, as all vocabulary items that were employed in the test, were familiar to participants.

As a result of the first stage of summarizing, three spreadsheets were created – one for each test. Each spreadsheet contained all relevant inflections produced by the testees within the phrases; each inflection tagged by each of eight tags. See a print screen in Appendix M illustrating a section of the initial Test 2 spreadsheet.

Overall, the first stage of data summarizing process, though having consumed considerably more time than it had been anticipated, has provided a detailed account of information about all relevant instances of inflection produced by participants during the three rounds of testing.

To prepare the recorded data for quantitative analysis, during the second stage of the summarizing process, I converted all tagged information about case inflections into statistical values by using Microsoft Access, which allowed me to do a number of queries. This allowed

me to summarize the raw scores for various categories, required for me to answer my Research Questions, namely, for each test overall, for each case, gender and case context, as well as for nouns with different degree of familiarity; these are presented in subsequent sections. Finally, the accuracy rates were calculated for each category and the tables were then fed into SPSS to perform statistical analysis.

6.2. General statistics for the three tests.

To answer the majority of my RQs (RQ1 – RQ6), I used the data collected during the post-testing (Test 2), while Test 1 provided the initial data for accuracy for Prep. (RQ2 and RQ5); and Test 3 (delayed) contributed some additional data for the first four RQs, regarding the retainment of accuracy. For this reason, in this section, I will present general statistics for the three tests and, in subsequent sections, I will focus my analysis on case inflection accuracy, as well as different potentially confounding variables in the data collected, to answer each of the RQs separately.

6.2.1. General statistics for Test 1

Test 1, “The Guessing Game” (see 5.4.3.), tested the production of nominal Prep. case forms in participants’ unprepared speech in isolation from Acc. It was recorded either at the end of Week 15 or at the start of Week 16 of the instruction period, before the Acc. was taught.

Altogether, in Test 1, 28 participants produced 434 inflection contexts, all for Prep. case in location context (which was the only context tested) with the mean of 15.5 inflections elicited per sample; with the lowest of 8 (P2013) and the highest of 21 case forms in a sample (P2006). The difference was due to how participants fulfilled the conditions of the Guessing Game. Some might have guessed the items’ locations too quickly (thus, their partners asked very few questions and produced limited number of inflections), while the others might have not been successful in guessing and enabled their partners to ask many questions to get the item’s location. This was controlled as much as it was possible during the test.

6.2.2. General statistics of the Test 2 data

Speaking Test 2 investigated the accuracy of the production of Prep. and that of Acc., as well as the overal accuracy, thus answering the first four of my RQs. In addition, to answer

RQ5, the effect of potentially confounding variables, such as gender, case contexts and familiarity of lexis, on case inflection accuracy was analysed (see 6.1.3).

Test 2 was administered to all experimental participants who completed the instruction period - 27 altogether (see 5.2.1). They took Test 2 online during the two weeks that followed the last learning session.

The total of 1198 relevant case inflection contexts were recorded initially, that is, for Prep. and Acc. cases. However, as the use of Acc. in the object case context was explained to participants during the last session of the instrunction period. Thus, the conditions for retrieval were different from the other contexts and did not comply with the principle of revisitting of the proposed Teaching Framework, making the inflection production in object context not representitive, with the strong potential of skewing the results. For that reason, 66 nouns (5%) in object context were taken out of the overall count, reducing the obligatory context to 1132, which is used in the rest of the analysis.

On average, participants produced 42 inflections per sample, with the smallest sample of 31 case forms (P2019) and the two largest samples containing 51 case suffixes each (P2025, P2029). The distribution for the entier set is normal.

Test 2 consisted of two parts - an Interview part and a Comics Task (see 5.4.4). According to my test design, the Interview part (Part 1) lasted 5 minutes and was much shorter than the Comics part (Part 2). Logically, the number of inflections produced in the Interview is significantly smaller – 318 inflections, which is 27% of the total inflections.

To investigate the accuracy of Prep. and Acc. (RQ2 and RQ3), as well as the differences in their production (RQ 4), the Test 2 data was devided into two case data sets. Participants produced more Prep. inflection contexts, 717 (60%), than those for Acc. – 481 (40%). The ratio for the number of inflections for Prep. and Acc. were exactly the same in each part of the test (the Interview and the Comics Test) - 60% : 40% respectively. The difference in the amount of production for the two cases was anticipated (see 5.4.4 for discussion) and will be discussed further in 6.7. The distribution of data for both case data sets was normal. The statistical analysis was performed separately on each case data set and the success rates were calculated from the total for the set.

Furthermore, to answer RQ5, the Test 2 data was also analysed for accuracy in different case contexts, as well as for the effects of gender and familiarity of lexis, which is presented in the subsequent sections and contributes more data to the analysis of each of the cases.

6.2.3. General statistics of Test 3 (delayed) and the comparability with Test 2

The delayed test (Test 3) was administered 6-8 weeks after the post-test (Test 2) (see 5.4.5), in order to establish, how well participants retained the skills of inflecting Russian nouns for case (specifically, Prep. and Acc.) in their speech. These data contribute to answering RQ1, RQ2 and RQ3, investigating the effectiveness of the proposed teaching framework for the production of the two cases, as well as partially to RQ5 and RQ6, dedicated to some factors affecting case inflection production.

Similarly to the first two tests (see 6.2 and 6.3), the Test 3 data were analysed for distribution of success rates for each case and each gender, as well as for each individual suffix. Due to the aims of the delayed testing, as well as because of the considerations of the scope of the current study, the case contexts and familiarity of lexis were not analysed for Test 3.

From 27 participants who took the post-test, 14 subjects took the delayed test (Test 3), which is just over a half. From the 17 high scorers in Test 2 (who had the success rate over 80% and constituted 62% of the 27 participants), seven took Test 3 (P2002, P2007, P2008, P2013, P2023, P2027 and P2029) and made exactly 50% of the 14 Test 3 participants. Thus, the balance between those who scored above average and below average in Test 3, was slightly in Favour of Test 2. In addition, it was checked that, from 10 participants with WM span above average for the group (that is, over 51 – see 6.2.3), only four took Test 3, again, creating an advantage for Test 2. That is why, I believe that it would be fair to say that the group of participants who took Test 3 was a good representation of Test 2 group, with slightly more participants at the lower end of the scale.

These participants produced the total of 605 relevant case inflection contexts for Prep. and Acc., which were considered obligatory contexts. This number could be approximated to a half of the inflection total in Test 2 (1132), which is appropriate for the lower number of participants in Test 3. On average, 43 case inflections were produced per sample, which is very similar to the average for Test 2 (42 inflections per sample), which confirms the comparability of the two tests. The largest samples (P2021, P2027, P2029) contained over 50 inflections, with the smallest sample (P2018) of 33 case suffixes, which is, again, in line

with Test 2. Interestingly, participant P2029 produced one of the largest samples in each of the tests.

The ratio between Prep. and Acc. cases in Test 3 (57% : 43% respectively), is similar to that in Test 2 (60% : 40%). Similarly to Test 2, more Prep. inflections were produced. The distribution of the data for both cases appears normal. In addition, the percentage of the Interview inflections in Test 3 (31%) is in line with that in Test 2 (27%) and was determined by my research design. Finally, the distribution of Prep. and Acc. inflection between the two parts of Test 3, namely, the Interview and the Comics Test, (55% : 45% respectively) also appears similar to that in Test 2 (60% : 40%). Considering the spontaneity of oral production, these ratios are very difficult to control and these figures can be recognized as comparable.

Overall, the general statistics of Test 3 demonstrates that the delayed test data (Test 3) can be accepted as a valid comparison for the post-test (Test 2), and the robustness of the test designs can be confirmed.

6.3. Results for the RQ1 – overall accuracy of case inflection production

To enable me to answer RQ1 investigating the effectiveness of the application of the proposed spiralling framework to teaching of Russian case inflection, the data from the post-test (Speaking Test 2) was used.

Overall, from the total of 1132 inflections, 27 participants produced 911 case forms, which were recorded as correct (see definition of “correct” in 6.1.3), with the mean of 33.74 and the median of 33. This constitutes the overall success of 80% which can be considered as acquired in SLA (e.g., Vainikka & Young-Scholten, 1994), as well as equating to a first grade in HE. This figure alone allows me to answer the RQ1 in the affirmative (see discussion in 7.1), demonstrating the effectiveness of the proposed Teaching Framework for teaching Russian case inflection.

The maximum score was 95% (P2008) and the minimum was 52% (P2019); these are further discussed in 7.4.2, together with other individual differences.

Furthermore, the data show that the success rates were slightly different for different parts of Test 2 (see Table 17 below).

Table 17

Success Rates for Test 2 Overall, for Its Two Parts, Compared to That for Familiar Items

	Total for Test 2	In Interview (Part 1)	In Comics task (Part 2)	Familiar
Number of correct inflections	911	262	649	634
Success rate	80%	86%	78%	82%

The table above demonstrates that the testees were more successful, with regard to case inflection, in the Interview part than in the Comics task (which contained 19% of unfamiliar nouns in case contexts) – 86% and 78% respectively. Also, the accuracy in the Interview (86%) is higher than that for familiar items for the entire Test 2 (82%) (see discussion in 7.1).

Further statistical analysis of Test 2 data was performed to investigate participants' case inflection production at different levels with more factors investigated. The results of these analyses are presented in the following sections.

Finally, delayed testing (Test 3 – see 6.2.3) was performed, in order to examine whether the production of case inflection would be retained at the same level over time. From the total of 605 case inflections recorded in Test 3, 14 participants produced 470 correct inflections, which constitute 78% success rate (see Table 18).

Table 18

Overall Raw Scores and Success Rates for Test 2 and Test 3.

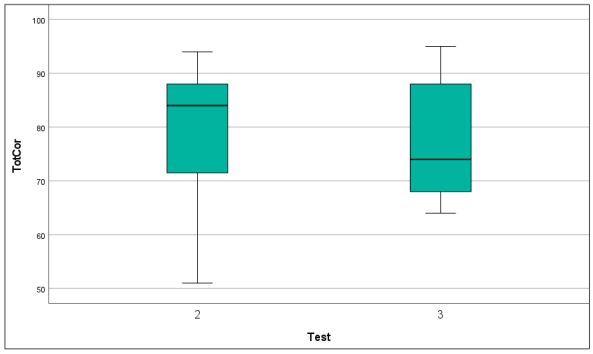
	Total inflections	Total correct inflections	Success rate
Test 2	1132	911	80%
Test 3	605	470	78%

As it can be seen from the table above, the success rate for Test 3 is very similar to the 80% overall success rate for Test 2, with 2% difference, which is not significant.

Furthermore, the boxplot in Figure 15 illustrates a change, which was not reflected by the overall scores – though the interquartile spread of data is similar in the two tests, the range of the Test 3 data at the low end is considerably reduced. The lowest score in Test 2 is 52% (P2019), while, in Test 3, it is 64% (P2028), though the median is marginally lower in Test 3. This means that the participants' performance has become more consistent.

Figure 15

Boxplot Illustrating Total Correct Scores in Test 2 (2) and Test 3 (3)



Considering the above analysis, it would be fair to recognise that the case inflection production skills, acquired by participants during the teaching intervention, were retained after 6-8 weeks, thus confirming the positive outcome of this part of my investigations, related to RQ1.

6.4. Results for RQ2 – accuracy for Prep. case inflection (initial statistics)

The accuracy for Prep. case inflection production (RQ2) was investigated in Test 1 and Test 2 (post-test), as well as in Test 3 (delayed) – see test aims and design in 5.4.3, 5.4.4 and 5.4.5. In short, the idea was to see how the case inflection accuracy rates reflect the increase in demand for participants' cognitive resources. To start with, in Test 1, Prep. infection was produced without interference of another oblique case (Acc.). This allowed learners to concentrate on Constituent Assembly, while Function Assignment was fairly straightforward. Then, in Test 2, Prep. had to be processed at the same time as Acc., and Function Assignment became a real challenge, while Constituent Assembly remained seemingly the same. This section present the initial statistics for Prep., without potentially confounding variables, namely, gender, case contexts and familiarity of lexis. The analysis of these will be carried out in the subsequent sections (to answer RQ5) and will contribute further data to RQ2.

Test 1 was administered when Prep. was the only oblique case that was taught and tested (see 6.2.1 for general Test 1 statistics).

In Test 1, from 434 nominal forms used in Prep. context, 367 inflections were marked as correct, constituting 85% success rate for the Prep. case inflection production. Therefore, at this stage, the answer to RQ2 can be affirmative, and the proposed spiralling framework could be provisionally considered as being effective for teaching Prep. case to complete beginners in classroom environment.

In Test 2, the Prep. case inflection accuracy was examined under different condition, as the Acc. case had been introduced into production, thus Test 2 data on Prep. further contributes to answering RQ2.

From the total of 717 case forms produced in Prep. obligatory contexts in Test 2, 575 were marked as correct, with the success rate of 80%. Using the SLA criteria of successful acquisition, Prep. case can be deemed as acquired, thus positively answering RQ2.

However, the scores for Prep. in Test 2 are lower than the rate in Test 1 (85%). The difference between Test 1 and Test 2 was analysed statistically. As the results for both tests were collected from the same participants, paired T-test was carried out. Because the results were obtained with different tests, the percentages of correct inflections were used. The difference between the two tests was negative (-5%), but not significant (p = 0.109). This instigated further investigations examining the factors influencing case inflection production (see RQ5). Thus, the subsequent sections also contribute to answering the RQ2 and the phenomenon of the drop of the initial success rates for Prep. is discussed in 7.3.3.

Furthermore, the success of Prep. case inflection production was examined in Test 3 (delayed). The success rate for Prep. in Test 3 was 77% (268 correct from the Prep. total of 347), which is 3% lower than in Test 2 but is exactly the same as the overall success rate for Test 3. The data for Prep. in the three tests is shown in Table 19.

Table 19

Prep. Case Inflection Production for the Three Tests.

	Total inflections for Prep.	Total correct inflections for Prep.	Success rate for Prep.
Test 1	434	367	85%
Test 2	717	575	80%
Test 3	347	268	77%

Though the success in production of Prep. case inflection dropped between Test 1 and Test 2 (possible reasons are investigated further), and the success rate went down slightly between the post-test (Test 2) and the delayed test (Test 3), none of the differences are significant. Also, even the lowest figure of 77% is way above the acceptable acquisition level of 60% for SLA, thus confirming the affirmative answer to RQ2. Finally, the participants' performance on Prep. will also be compared to that on Acc. in 6.6.

6.5. Results for RQ3 – accuracy for Acc. (initial statistics)

To answer the RQ3, the data from Acc. data set from Test 2 was used (see 6.2.2 for Test 2 general statistics). It is worth noting that Acc. was taught during the last five weeks of the Instruction period and, therefore, was not tested in Test 1 (see research design in 5.3.2 and 5.4.3).

From the total of 481 obligatory contexts for Acc. in Test 2, participants produced 335 correct case inflections, which renders 80% success rate. Similarly to Prep. (see 6.4.2), this is considered an expected level of performance in both SLA and HE (see 6.3), thus affirmatively answering the RQ3.

The delayed testing also confirmed the successful acquisition of Acc. case inflection by the teaching intervention participants, as from 258 nouns in the obligatory Acc. context, 202 were inflected correctly. The success rate was calculated at 78%, which, similarly to Prep., is well above the acceptable acquisition level, and only 2% below the scores for Test 2 (see Table 20). This shows that participants retained the skills for inflecting nouns for Acc., acquired during the Instruction period, thus supporting the positive answer to RQ3.

Table 20

Acc. Case Inflection Production for Test 2 and Test 3.

	Total inflections for Acc.	Total correct inflections for Acc.	Success rate for Acc.
Test 2	481	335	80%
Test 3	258	202	78%

However, in the course of my investigations, when the other potentially confounding variables are factored into analysis, it has become apparent that the case inflection production, for both Prep. and Acc., appears to be a considerably more complex process.

Therefore, both the Prep. and the Acc. data is examined again in the subsequent sections, to answer RQ4 and RQ5.

6.6. Results for RQ4 - the difference in participants' performance on Prep. and Acc. (initial statistics)

The analysis in the two previous sections clearly shows that the accuracy rates for Prep. and Acc. inflection in the post-test (Test 2) were exactly the same and in the delayed test (Test 3) were almost the same – with 1% difference. This suggests that participants performed very similarly in the two tests where both cases were in production. In Test 1, the accuracy for Prep. (which was the only case tested) was slightly higher, but the difference still was not significant (see Tables 14 and 15). Therefore, at this stage of analyses, the answer to the RQ4 appears to be negative. However, in order to observe the dynamics of changes in the case inflection data across the three tests, I created the boxplots for each case within each test (see Figures 17 and 18).

Figure 16

Success Rates for Prep. in Test 1 (1), Test 2 (2) and Test 3 (3)

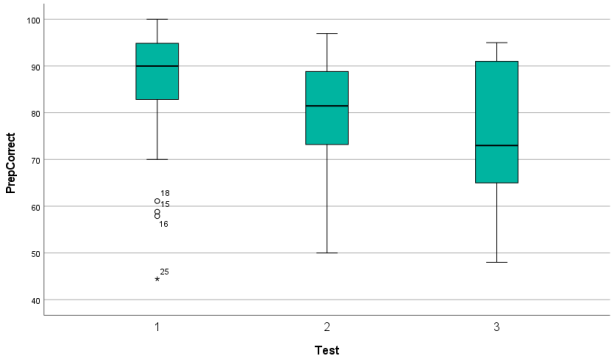
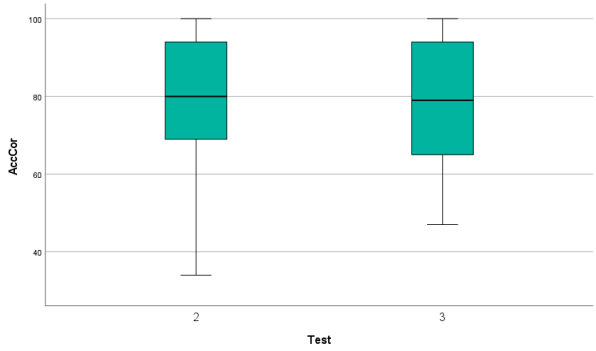


Figure 17

Success Rates for Acc. in Test 2 (2) and Test 3 (3)



The boxplots above show that the production of case inflection for Prep. and Acc. develops in rather different ways. For Acc. (which was not taught before Test 1), the interquartile range and medians are very similar in the two tests, while the overall range of data is reduced, pointing to more stable performance by participants. In contrast, for Prep., the spread of data increases throughout the tests, and the medians go down. However, the outliers, which are present in Test 1, become part of an overall range of scores and the number of scores above the median increases. Though these are also the signs of the group performance becoming more consistent, the developments in the production of the two cases appear very different and, therefore, were investigated further.

6.7. Results for RQ5 – analysis of case inflection production with gender as a potential confounding variable

RQ5 was posited to investigate the three factors that could potentially affect the acquisition of case inflection, namely, gender, case contexts and familiarity of lexis (see 5.4.1). In this section, I will present the statistical results related to the role of grammatical gender in the success of the production of Russian case inflection. The other two factors will be analysed in the following two sections.

6.7.1. Gender statistics for Test 1

As some differences in case inflection success rates were expected between masculine and feminine case forms, due to an extra step in Constituent Assembly in feminine, despite the case suffix being the same (see 4.1.2. and 4.2.2.), the data for all three tests was further analysed with grammatical gender as a confounding variable.

From the total of inflections in Test 1, there were 206 masculine case forms and 210 feminine case forms (see Table 21 below). The balance between masculine and feminine was mainly dictated by the choice of items used in the test (see 5.4.3 and Appendix F). In addition, 18 instances of neuter were registered (*âbloko* for “an apple” and *moloko* for “milk”). Within the Test 1 design, these were intended for the use in a Subject function, as pragmatically, they are unlikely locations, but during the Guessing Game (Test 1), some of the participants experimented with these (also, see reasoning for not investigating neuter in 5.4.1). Half of the participants (14 participants) did not use neuter in location contexts at all. It is apparent that the neuter sample is very small (18 case forms from the total of 434, which makes 4%) and cannot be considered representative. However, some neuter examples are used for the qualitative analysis.

The distribution of correctly produced inflection between genders is presented in Table 21.

Table 21

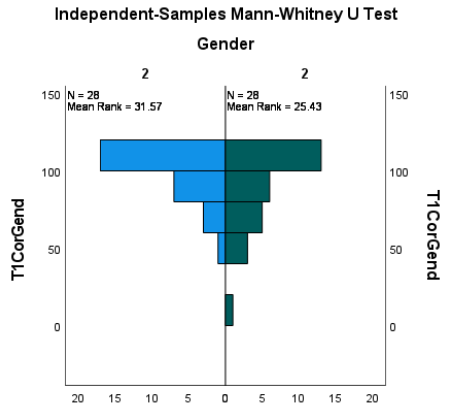
Prep. Inflection Success Rates for Masculine and Feminine in Test 1

	Total inflections	Total correct	% from total for gender
Masculine	206	184	91%
Feminine	210	177	82%

As it can be seen from the table above, the success rate for masculine was 9% higher than that for feminine (91% and 82% respectively). Therefore, the difference in accuracy was analysed statistically. As the data was negatively skewed, the parametric Independent Sample test was not suitable, and the non-parametric Mann-Whitney U Test was chosen. The statistics analysis showed that the difference between genders in Test 1 was not significant, as the calculated p-value of 0.125 is higher than significant 0.05. The graph in Figure 18 illustrates that, despite some differences, the data patterns for each gender appear to be similar.

Figure 18

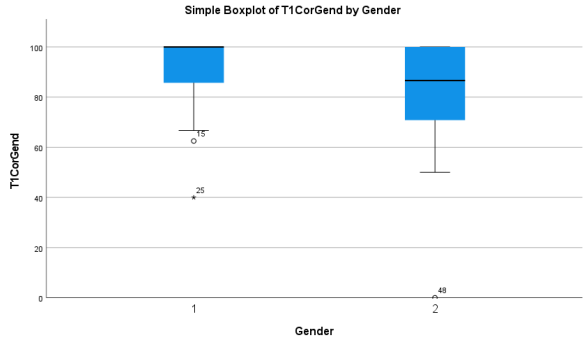
Graph Illustrating the Distribution of Correct Prep. Case Forms Between Genders



However, in order to see, how the correct scores are distributed within each gender, the boxplots were created for masculine and feminine case inflection scores (see Figure 19).

Figure 19

Boxplot Illustrating the Spread of Test 1 Prep. Success Rates for Masculine (1) and Feminine (2)



The graph above makes it clear that the lower mean is not the only difference between the genders, as the feminine data appears to be considerably more spread. That is why, the role of gender in inflection production was investigated further in Test 2.

Moreover, when the number of incorrect inflection for masculine and feminine case forms was analysed separately, the difference of total incorrect between genders was very significant, with *p*-value of 0.000000036. As the percentage of incorrect inflection totals in Test 1 was fairly low (15% from the total inflections), this is considered indicative, rather than statistically viable.

Furthermore, the two types of errors, were analysed for different gender, namely, base form errors and postponed inflection errors (see Tag 5 in 6.1.3 for definitions). P2026 sample in Example 1 from the "Guessing game" test (Test 1) shows both types of errors for feminine case forms, as well as the emergence of these errors. This will also be used for qualitative analysis in 6.11.2.

Example 1. Prep. masculine and feminine forms with two types of errors.

P2012: *Âbloko – v paket**e**?* P2026: *Net*	P2012: Is the apple in the plastic bag*[Prep.masc.correct]*? P2026: No.
P2012: *Âbloko – na sok**e**?* P2026: *Net*.	P2012: Is the apple on the (carton of) juice*[Prep.masc. correct]*? P2026: No.
P2012: *Âbloko – v stol**e**?* P2026: *Net.*	P2012: Is the apple inside the desk*[Prep.masc.correct]*? P2026: No.
P2012: *Gde âbloko?* P2026: *Âbloko – na čašk**a**.* P2012: *Âbloko – na čaške.* *Swap.*	P2012: Where is the apple? P2026: The apple is on the cup*[Prep.fem.base form error].* P2012: The apple is on the cup*[Prep.fem.correct].*
P2026: *Moloko. Moloko – na korobk**a**?* P2012: *Net*.	P2026: Milk. Is the milk on the box*[Prep.fem.base form error]*? P2012: No.
P2026: *Moloko – v korobk**a** ...k**e**?* P2012: *Net*.	P2026: Is the milk in the box*[Prep.fem.postponed inflection error]*? P2012: No.
P2026: *Moloko – v paket**e**?* P2012: *Net*.	P2026: Is the milk in the plastic bag*[Prep.masc.correct]*? P2012: No.
P2026: *Moloko – na korobk**e**?* P2012: *Net*.	P2026: Is the milk on the box*[Prep.fem.correct]*? P2012: No.

Table 22 presents the distribution of these two types of errors between masculine and feminine.

Table 22

Total Incorrect and Error Type Distribution Between Genders

	Total incorrect	Base form errors	Postponed inflection errors
Total	54	34	20
Masc.	21	13	8
Fem.	33	21	12

It is evident that more base form errors were recorded than postponed inflection errors (34 and 20, respectively). However, participants produced more of both types of errors for feminine than for masculine case forms.

Similarly, the data on attempts were also summarized for the two genders separately and are presented in Table 23. They show that participants produced a very few attempts in Test 1 – in 47 inflection contexts from 434 of total inflections. Nevertheless, it is apparent that the absolute majority of these are base form attempts – 38 instances from 47 attempts produced. Moreover, 5.6 times as many base form attempts were made on feminine locations than on masculine (21 and 5 instances respectively). It is also important to note that the 33 of these attempts (70%) resulted in correct inflection, with only four rendering a base form error.

Table 23

Types of Attempts and Their Distribution Between Genders

	Total attempts	base form attempts	postponed inflection attempts
Total	29	28	5
Masculine	5	5	0
Feminine	22	21	5

Finally, there is a very small number of postponed inflection attempts, in comparison with the postponed inflection recorded as errors, 5 and 20 respectively (see Tables 22 and 23). However, all of the postponed inflection attempts are produced on feminine case forms. Due to small numbers of errors and attempts, the above results cannot be statistically viable but

they are definitely indicative of the differences in Russian case inflection production between masculine and feminine; they will be discussed in 7.2.1.

Overall, though none of the differences in the production of case inflection for masculine and feminine can be considered statistically significant in Test 1, they illustrate some tendencies in inflection processing and have prompted further investigations in Test 2 and Test 3.

6.7.2. Gender statistics for Test 2

Following the findings of Test 1, regarding the differences in the case inflection production between genders, the Test 2 data was also analyzed from the point of view of gender distribution (see the set of inflections for Test 2 in Table 16).

As the production in Test 2 was more varied than in Test 1, controlling the balance between the number of masculine and feminine case forms was considerably more challenging, especially in the Interview part of the test. That is why the distribution of genders is uneven (see Table 24).

Table 24

Case Inflection Production Distribution Between Masculine and Feminine

	Total inflections	Prep. inflections	Acc. inflections	Inflections in Interview
Masc	783	515	270	251
Percentage from above	69% from total inflections	70% from total Prep	65% from total Acc	83% from total Interview
Fem	349	202	145	51
Percentage from above	31% from total inflections	30% from total Prep	35% from total Acc	17% from total Interview

The table above demonstrates that participants produced more masculine case forms - 783 (69% of the total for Test 2), as opposed to 349 feminine (31%). The ratio of masculine and feminine is similar for each of the cases – 70% : 30% for Prep. and 65% : 35% for Acc. However, the striking difference between masculine and feminine sets is that considerably more masculine case forms are produced in the Interview part – 83%, while the feminine case inflections in Part 1 constitute only 17%. Interestingly, when case was factored in the distribution of genders for the Interview, it appeared that 50% of all Interview inflections are masculine Prep. – 158 from the total of 318 for the Interview (see discussion in 7.3.3)

In order to enable me to compare the accuracy of the case inflection production between different cases and genders, as well as between individual suffixes, the success rates were calculated as percentages from the totals for each case and for each gender, as well as for each of the four suffixes. The distribution for different data sets for Test 2 varied, therefore, for statistical analysis, a range of parametric and non-parametric tests was used.

First, the success rates for masculine and feminine case inflection were calculated (see Table 25).

Table 25

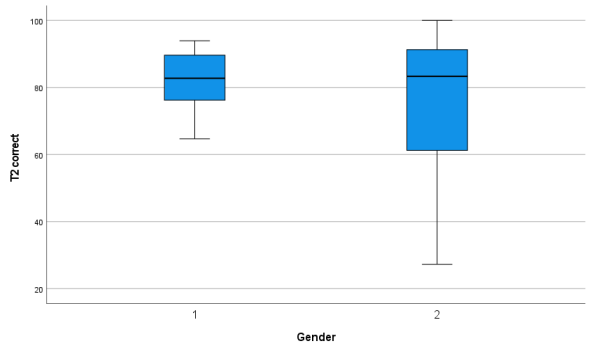
Case Inflection Success Rates for Masculine and Feminine for Test 2

	Masculine	Feminine
Total correct for gender	644	261
Success rate	82%	74%

According to the table above, the case inflection production success rate is higher for masculine than for feminine – 82% and 74% respectively. To investigate these differences, statistical tests were performed.

To start with, the distribution of correct inflection for each gender was checked and appeared normal. That allowed me to carry out parametric testing on these data. The total correct scores for each gender were compared, using independent T-test, with percentages of correct inflection as dependent variable and gender as factor. Though the difference between success rates for masculine and feminine was not significant (p-value of 0.09), there was a considerable difference in variance (87.92 and 403.71 respectively), which is clearly demonstrated by the boxplot in Figure 20. The graph shows that, though the medians for masculine (1) and feminine (2) are similar, the feminine data are more spread, also having a larger share of lower scores.

Figure 20
Boxplot Illustrating the Spread of Case Inflection Scores for Masculine (1) and Feminine (2) in Test 2



These differences in the success rates for each gender had the potential to stem from the difference in processing of different case suffixes (as Acc. suffixes were different for each gender), thus directly contributing to RQ5. For that reason, the percentages of correct inflections for Prep. and Acc. were analysed separately for masculine and feminine.

Table 26

Prep. and Acc. Success Rates for Each Gender (Test 2).

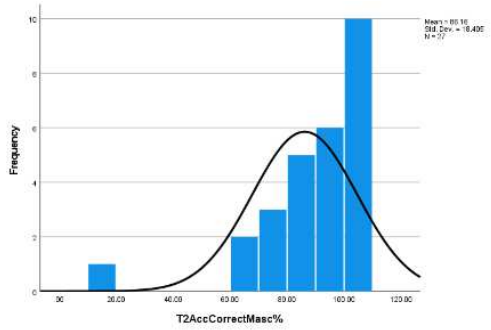
Gender		Prep.	Acc.
Masc.	Raw correct scores	415	232
	Percentages of correct from total for masculine	80%	86%
Fem.	Raw correct scores	162	105
	Percentages of correct from total for feminine	80%	72%

Table 26 shows that, on one hand, the accuracy for masculine and feminine was exactly the same in Prep. (80%), while, on the other hand, the percentages of correct case forms were rather different in Acc. (86% for masculine and 72% for feminine). This demonstrates that learners performed differently for different cases, thus positively answering RQ4, which is contrary to the initial conclusion made in 6.6.

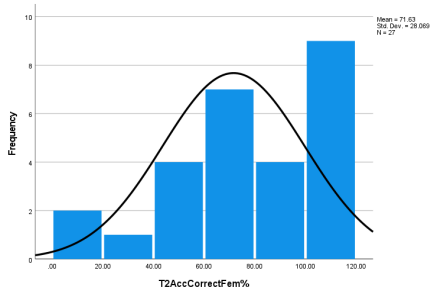
The histograms showed that the Prep. data for the two genders were normally distributed, while the Acc. data was negatively skewed for both masculine and feminine (see Figures 22 and 23). That meant that there were more Acc. scores produced at the lower end of the scale with the sharp rise to higher accuracy.

Figure 21 **Figure 22**

Distribution of Correct Masculine Case Forms in Acc. (Test 2)



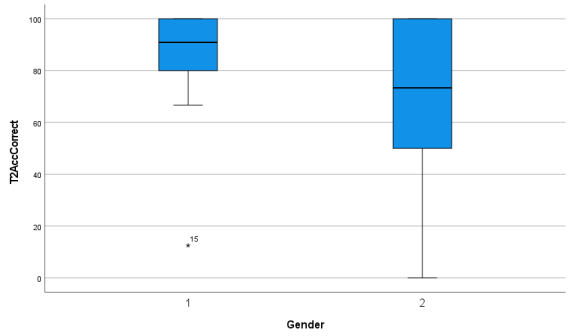
Distribution of Correct Feminine Case Forms in Acc. (Test 2)



As the assumption of normality for these data had not been met, the use of parametric tests was not suitable, and non-parametric Mann-Whitney U test for independent samples was performed instead. The difference in correct production between masculine (1) and feminine (2) for Acc. was significant (p = 0.049). There was a considerable difference in Mean Ranks in Acc. – 31.61 : 23.39. Also, the boxplots for the accuracy for two genders in Acc. show the spread of the data, which is considerably larger for feminine (see Figure 23).

Figure 23

Boxplots Illustrating the Spread of Data Between Genders in the Acc.



Furthermore, in order to investigate the drop in Prep. accuracy (see 6.4), the differences in Prep inflection production for masculine and feminine were compared between Test 1 and Test 2 (see Table 27).

Table 27

Prep. Case Success Rates for Masculine and Feminine in Test 1 and Test 2

		Test 1	Test 2
Prep.	Masculine	91%	80%
	Feminine	82%	80%

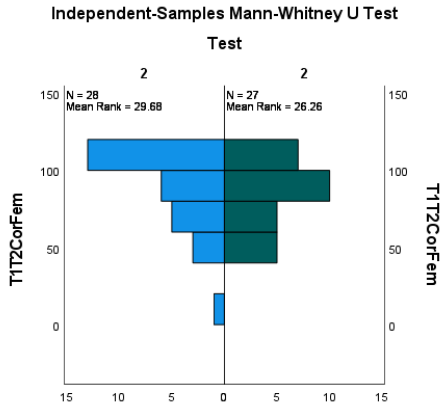
As Test 1 data for each gender were not normally distributed, the non-parametric option for comparing genders effects in the two tests statistically was chosen, namely Mann-Whitney U Test. The difference was significant only for the masculine - 91% vs 80%, with Asymptotic Sig. < 0.001, as opposed to feminine – 82% vs 80%, with Asymptotic Sig. = 0.417. The graph in Figure 24 clearly shows that the distribution patterns in the two tests are rather different for masculine, where the Test 1 data is negatively skewed and Test 2 Prep. scores display normal distribution. Figure 25 demonstrates that the Prep. feminine data patterns are quite similar in both tests.

Figure 24

The Difference Between Correct Scores for Prep. Masculine in Test 1 and Test 2

Figure 25

The Difference Between Correct Scores for Prep. Feminine in Test 1 and Test 2



The significant difference between Prep. masculine success rates for the two tests, as well as well as very different distribution of scores suggest that the drop in Prep. accuracy between Test 1 and Test 2 occurred only for masculine case forms.

6.7.3. Gender statistics for Test 3

In delayed testing data (Test 3), the split of the produced case inflection between genders (69% masculine case forms : 31% feminine), though varying slightly between cases, is in

line with figures for Test 2 (Table 24 in 6.7.2). The case inflection success rates for masculine and feminine, produced in the post-test and in the delayed test are summarized in Table 28.

Table 28

Case Inflection Success Rates for Masculine and Feminine in Test 2 and Test 3

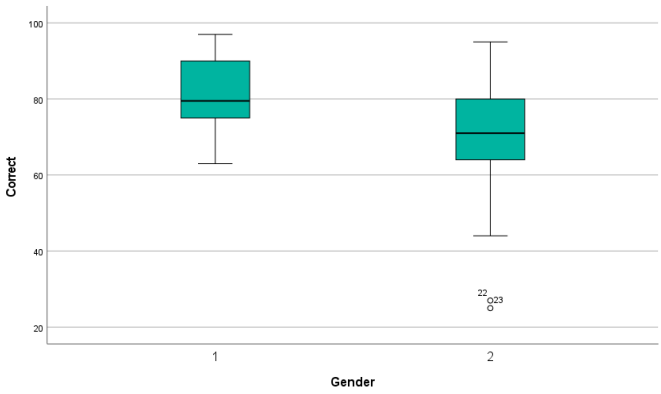
Gender		Total correct for gender Test 2	Total correct for gender Test 3
Masc.	Raw scores	644	341
	Percentages of correct from total for gender	82%	81%
Fem.	Raw scores	261	129
	Percentages of correct from total for gender	74%	70%

The table above clearly shows that the percentage of correct masculine case forms in Test 3 (81%) is very close to both overall success rate for Test 3 (78%) and to Test 2 masculine scores (82%). Participants' accuracy for feminine in Test 3 (70%) also does not appear considerably different from Test 2 feminine scores (74%). However, it is 8% lower than the overall success rate in Test 3 (78%). This also creates a larger gap (11%) between masculine and feminine scores (81% : 70% respectively) than that for Test 2 (82% : 74% respectively), which was 8% and was not significant (see Table 25). Therefore the Test 3 difference between masculine and feminine scores was tested statistically by running an Independent Sample t-test. As the distribution appeared normal, the parametric test was suitable. Even the lowest p-value of 0.055 falls short of significant, demonstrating that the difference between genders was not significant. Nevertheless, the increased difference in accuracy between genders has prompted me to investigate this further.

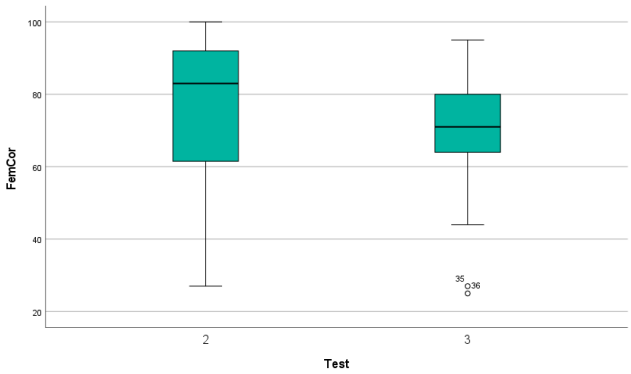
To start with, to see whether the spread and the range of data for each gender changed in Test 3 (in comparison with Test 2), the boxplots were created (see Figure 26). Visually, the Test 3 boxplots appear much more similar to each other than those for Test 2 (see Figure 20), as the range of scores is much reduced in feminine, though the mean for feminine is slightly lower than that for masculine. This change is vividly illustrated in Figure 27.

Figure 26. **Figure 27.**

Boxplots Illustrating the Spread of Data for Masculine and Feminine Case Inflections in Test 3



Boxplots Illustrating the Spread of Data in Feminine Case Inflections for Test 2 and Test 3



However, there are two outliers at the low end in Test 3, who produced very low scores for feminine (P2017 – 27% and P2018 – 25%), despite good overall scores for Test 3 (73% and 67%, respectively).

Then, factoring gender in together with case resulted in the following success rates for – see Table 29.

Table 29

Prep. and Acc. Success Rates for Masculine and Feminine in Test 2 and Test 3.

Gender		Prep.		Acc.	
		Test 2	Test 3	Test 2	Test 3
Masc.	Raw correct scores	415	210	232	131
	Percentages of correct from total for the case within gender	80%	77%	86%	89%
Fem.	Raw correct scores	162	58	105	71
	Percentages of correct from total for the case within gender	80%	78%	72%	64%

Overall, none of the corresponding rates for Test 2 and Test 3 were found significantly different. From Table 29, it is clear that the scores for Prep. for the two genders in Test 3 are almost identical (77% and 78%) and are the same or nearly the same (with 1% difference for masculine), as the overall success rate for Test 3 (78%). Also, they are very similar to the Prep. scores for both genders for Test 2. In contrast to Prep., for Acc., the case accuracy for

masculine has increased (though only by 3%), while for feminine, the rates went down (by 8%), producing a larger gap. For this reason, the Acc. data were investigated further.

First, the difference between the accuracy for Acc. feminine in Test 2 and Test 3 was statistically tested. As the data in the Test 2 was not normally distributed, the non-parametric alternative to the paired t-test was chosen, namely, Wilcoxon Signed-rank test. The p-value of 0.551 suggests that the difference between the Test 2 and Test 3 for Acc. feminine is not significant and the histogram in Figure 28 illustrates the similarity of the data patterns for the two tests. The above confirms that there are no significant differences between the Test 2 and Test 3 accuracy for any of the inflections even when gender is factored in. This demonstrates that after 6-8 weeks from the end of the instruction period, the participants retained the skills that they have acquired during the teaching intervention, thus further supporting the affirmative answers to RQ2 and RQ3.

Figure 28
Feminine Acc. Success Rates for Test 2 and Test 3

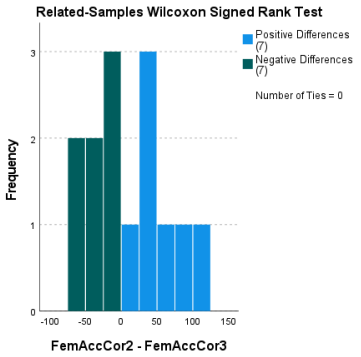
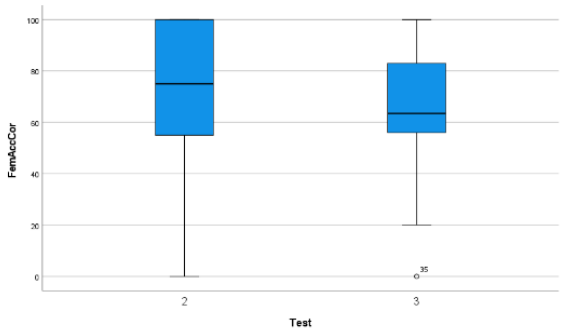


Figure 29
Boxplots Illustrating the Spread for Acc. Feminine Scores for Test 2 and Test 3



At the same time, the boxplots in Figure 29 (see above), show that the spread of feminine Acc. correct inflection was reduced and the range got smaller in Test 3, so the Test 3 boxplot appears more symmetrical. This is similar to the differences between overall feminine sets in both tests (see Figure 27).

The main result for Test 3 was that there were no significant differences between case inflection success rates in the post-test and in the delayed test, for any of the suffixes. In addition, Test 3 demonstrated, similarly to Test 2, that there are considerable differences in case inflection production between genders overall, as well as within each case. Thus, in

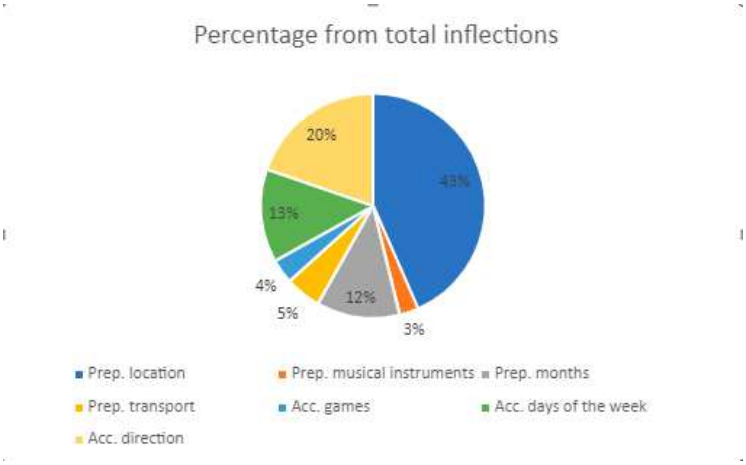
answer to the RQ5, gender can be considered one of the crucial factors affecting the case inflection production for Prep. and Acc.

6.8. Results for RQ5 - analysis of case inflection production with case contexts as a potential confounding variable

Another factor that was selected to be investigated in RQ5 as having the potential to affect the success of case inflection production, is case contexts (see discussion in 5.4.1). The pie-chart in Figure 30 shows the distribution of inflection production between case contexts for each of the cases. It might be worth noting that, in Test 1, only one case context was tested, namely, the Prep. location, while, in Test 3, case contexts were not investigated, due to the scope of the present study.

Figure 30

Distribution of Inflection Production Between Case Contexts in Test 2



It is apparent from the chart above that the amount of inflection produced by the participants varies a great deal between different case contexts. The Prep. location context stands out as providing the largest number of case forms in the test (491) which makes 43% of the overall number of inflections. It is clear that this case context is responsible for tipping the balance between Prep. and Acc. production in favour of Prep. (see 6.2.2). The second largest case context is Acc. direction (20%); it is also the largest for Acc.. Next, time references, namely, months (Prep.) and days of the week (Acc.), delivered similar percentages of inflections (12% and 13% respectively). The smallest number of inflections was produced for musical instruments (Prep.) and games (Acc.) – 3% for either of the contexts, with the remaining 5% supplied for transport case context. (Note that the object context was excluded from the

analysis, see 6.2.2.) It might also be worth mentioning that musical instruments and games contexts were only produced in the Interview part of the test (see test design in 5.4.4). These differences are discussed in 7.2.3.

As different case contexts varied in the number of items produced for the context (see Table 30), in the type of the lexical set that the nouns belong to (see definitions in 4.1.4), as well as the time of introduction during the instruction period, the accuracy was also calculated for each of the case contexts separately and are presented in Table 30. It is important to note that the number of items in different contexts vary, therefore the percentages were calculated from the total for the context.

Table 30

Success Rates for Case Contexts.

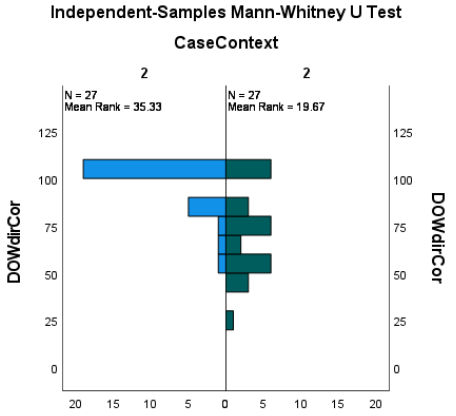
	Prep.				Acc.		
	location	musical instruments	months	transport	days of the week	games	direction
Correct case inflections	395	29	105	46	143	40	151
Percentage from total for case context	80%	93%	77%	79%	94%	100%	68%

As it can be seen from the above table, the highest percentages of correct inflections are supplied for games, days of the week and musical instruments (100%, 94% and 93% respectively), with location being next one down (80%). The lowest score is found for direction context (68%), with months and transport in the middle (77% and 79% respectively). This data already demonstrates that the accuracy for case inflection is different in different case contexts, thus contributing for RQ5.

The accuracy for only two contexts within one case was compared statistically, namely, days of the week and direction, as they had comparable number of inflections but very different success rates. As the data for the former was very skewed, the non-parametric test was performed. With p-value less than 0.001, the difference between the samples for the two contexts was very significant. The distribution patterns are also demonstrated by the graph in Figure 31.

Figure 31

Distribution of correct Acc. scores between two main Acc. case contexts, days of the week (DOW) and direction



To contribute further to answering RQ5, the relationship between the two factors, namely, gender and case contexts, was investigated and the distribution of the case contexts was analysed for each gender (see Table 31).

Table 31

Distribution of Genders Between Case Contexts.

		Prep.				Acc.		
		location	musical incstruments	months	transport	days of the week	games	direction
Masc.	Total	303	21	136	55	90	40	140
	% from total for the case for masc.	59%	4%	26%	11%	33%	15%	52%
Fem.	Total	190	10	X	3	63	X	83
	% from total for the case for fem.	95%	5%	X	1%	44%	X	56%

It is clear from the above that some case contexts require only or predominantly masculine items. This could not be addressed in the research design, as, for example, months are all masculine, constituting 26% of all masculine Prep. inflections. The small number of nouns for transport (55), which are mainly masculine, add another 11% to masculine Prep, leaving 59% to the location context. At the same time, feminine Prep. is made predominantly from location inflection – 95%, with musical instruments adding a tiny portion of 5%.

In the Acc., the distribution of inflections between case contexts is relatively more even. The feminine part of the direction context (56%) is balanced by the half of the recorded days-of-the-week inflections (with neuter *voskresen'e* (for “Sunday”) not counted, see 2.6.7), contributing 44% to the feminine part of the case data set. In addition, masculine inflection data set includes games (40), which added 15% to the masculine Acc. This data will be very important in the discussion of gender effects in 7.3.3.

Only the direction context was analysed on the subject of difference in accuracy between genders, as it had the lowest success rate (68%) among the case contexts (see Table 30). The raw scores and percentages for each gender in the direction context are presented in Table 32.

Table 32

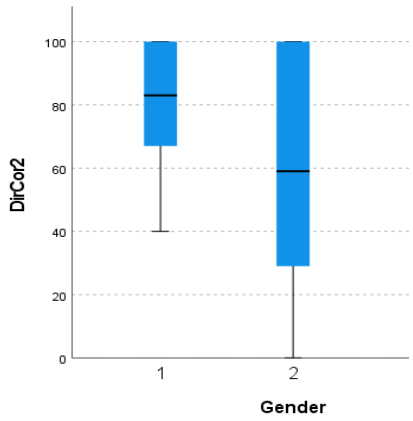
Raw Scores and Success Rates for genders in direction context (Acc.)

Gender	Total	Correct	Percentage from total for context
Masculine	140	105	75%
Feminine	81	46	57%

The difference in correct scores between masculine and feminine for the direction context appears larger than that for total Acc. (86% and 72% respectively; see Table 29). In order to find the significance of this difference, the means were compared. Though the distribution of the samples in histograms appeared to be approximately normal, the assumption of the equal variance was not met. Also, the spread of the correct inflection production for each gender was very different (see boxplot Figure 32) - the range of scores in feminine for this case context is the largest in Test 2 (from 0 to 100%).

Figure 32

Boxplots of Success Rates for Masculine and Feminine in Direction Context (Acc.)



Considering the above, it was decided to run a non-parametric Mann-Whitney U test. The acquired p-value of 0.027 has confirmed that the difference in the correct case inflection production for masculine and feminine in the direction case context, is significant. It is also more significant than the difference between genders for the Acc. case (p=0.049).

The above statistics in this section, confirms an important role of case contexts, in addition to the crucial role of gender, in Russian case inflection production, contributing to answering the RQ5. At the same time, the data above supports the affirmative answer to RQ4, given in 6.7, regarding differences in participants' performance on the two cases.

6.9. Results for RQ5 - analysis of case production with familiarity of lexis as a potential confounding variable

The next factor which was investigated for RQ5, was the familiarity of the items that were inflected, which was expected to increase or decrease the processing load and consequently, to affect the success of inflection production (see 3.3.2, 4.2.2 and 5.4.1 for discussion). Note that this confounding variable was only investigated in Test 2. In order to examine this aspect of inflection processing, the success rates were compared between different vocabulary sets, namely, familiar, rare and new (see 5.4.4 and Tag 8 in 6.1.3).

6.9.1. General statistics on familiarity of lexis in Test 2

Overall, 67% of nouns that were produced by the testees in obligatory case contexts, were classed as familiar, with almost identical percentages for the two cases (66% and 67%). Unlike in Test 1, in Test 2, in line with my research design, 18% of the vocabulary were new items that participants had never been exposed to before. The remaining 15% were classed as rare (see 6.1.3 for description). The number of new items for each of the two cases was nearly the same – 18% from the total for Prep. and 17% from the total for Acc.

It is logical that the absolute majority of lexical items used by the testees in their Interview part of the test were familiar. The only rare items which were recorded in the Interview were four musical instruments and three games, which were not practiced by the whole group during the instruction period (see Table 33 below).

Then, the split of the vocabulary for each category between case contexts was analysed (Table 33). This also reflects which lexical set the items belong to (see definitions in 4.1.4), as well as giving more insights for the discussion in 7.3.4.

Table 33

Distribution of Familiar, Rare and New Vocabulary Between Case Contexts

	Prep.			Acc.			
	Location	musical instruments	months	transport	days of the week	games	direction
Familiar	312	24	136	3	152	37	103
% from total for case context	64%	77%	100%	5%	100%	93%	46%
Rare	59	3	X	55	X	3	38
% from total for case context	11%	10%	0	95%	0	7%	17%
New	123	X	X	X	X	X	81
% from total for case context	25%	13%	0	0	0	0	37%

For obvious reasons, nouns belonging to close lexical sets, namely, months and days of the week, only figure as familiar. The majority of items in limited sets, musical instruments and games, were also familiar, 77% and 93% respectively. Very occasional rare items were recorded, when a testee used something which was not used in class, e.g., *truba* for “a trumpet” (P2008) and *kviddič* for a game of “Quidditch” (P2015). However, transport nouns, though belonging to limited lexical category, were mainly classed as rare, as these were introduced in session 19 and participants did not have time to use them as familiar. The largest amount of new vocabulary items were found in location and direction contexts, where unfamiliar Russian place names were used in the Comics task – 25% and 36% respectively.

6.9.2. Case inflection success rates for familiar, rare and new vocabulary

In order to answer RQ5 with regard to familiarity of lexis, the success rates were calculated separately for familiar, rare and new vocabulary. Table 34 displays the means for each of the set of the vocabulary, which clearly decrease as the familiarity of the vocabulary goes down – 79% for familiar, 72% for rare and 66% for new.

Table 34

Case Inflection Success Rates for Familiar, Rare and New Vocabulary.

Vocabulary set	Mean	Standard deviation
Familiar (1)	79.41	11.507
Rare (2)	72.30	22.160
New (3)	66.41	17.158

Though the distribution of each vocabulary set was normal or approximately normal, Standard Deviations varied (see Table 34). For that reason, non-parametric alternative to one-way ANOVA, namely Kruskall-Wallis Test, was chosen for statistical analysis. The Hypothesis Test Summary presented the p-value of 0.025, confirming that the difference between success rates for the three vocabulary sets was significant, thus rejecting the null hypothesis. However, the pairwise Comparisons in Table 35 indicate that Adjusted Sig. was below 0.05 only between familiar and new lexical sets, with p=0.02. This is also demonstrated by mean ranks nodes in Figure 33.

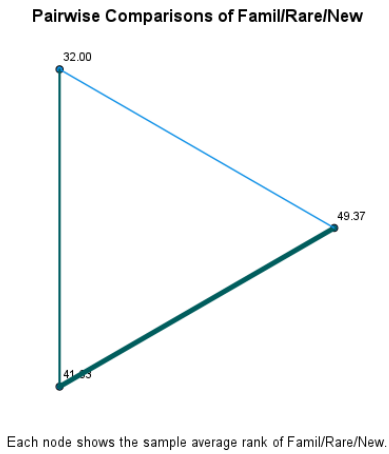
Table 35

Pairwise Comparisons Between Different Vocabulary Sets - Familiar (1), Rare (2), New (3)

Comparison of the three vocabulary sets	Test Statistic	Adj. Sig.[a]
3-2	9.630	0.396
3-1	17.370	0.020
2-1	7.741	0.678

Figure 33

Mean Ranks for Kruskall-Wallis Test with Familiarity of Lexis as a Variable



This result by itself demonstrates the significance of familiarity of lexis as a factor affecting case inflection accuracy, thus answering the RQ5 in the affirmative with regard to this line of investigation.

However, in addition, the case accuracy rates for familiar and new items were calculated separately for the two contexts, where the absolute majority of new items was used – location and direction. As the difference between case inflection accuracy rates was significantly different for the two genders within these contexts (see Table 32 and Figure 32 in 6.8), the effects of familiarity of lexis was also investigated for the direction contexts with gender as an additional variable (see Table 36).

Table 36

Case Inflection Success Rates for Familiar and New Vocabulary Within Location and Direction Contexts, with Gender Factored in

	Location context (Prep.)			Direction context (Acc.)		
	total for context	masculine	feminine	total for context	masculine	feminine
familiar	82%	82%	82%	65%	66%	63%
New	74%	74%	74%	70%	82%	50%

Table 36 demonstrates that the accuracy is higher for familiar items in location context (Prep.), and is consistent across the two genders for that case – 82% for familiar and 74% for new. For the direction context (Acc.) though, the success rates display great variation. Contrary to the overall result for the familiarity of lexis (see Table 34) and to the results for

the Prep. location context, the overall scores for the Acc. direction context are lower for familiar vocabulary and significantly lower than the percentages for familiar items in the location context. Moreover, though the overall score for new items in the direction context is only marginally lower than that for location, new items accuracy percentages calculated for masculine and feminine are significantly different, with $p<0.001$ (82% and 50% respectively). At the same time, the scores for familiar items in the direction context (Acc.) are fairly similar, with 1-3% differences.

The analyses in this section provides four main results. First and most important, familiarity of inflected nouns is an important factor affecting case inflection production; this contributing to RQ5. Second, though normally the increase in familiarity corresponds to the increase in accuracy, the familiarity of lexis can affect the success rates in a negative way too (see discussion in 7.3.4). Third, the fine-grained analyses above further supports the conclusion that, despite both Prep. and Acc. having absolutely identical overall success rates, participants' performance on the inflection production for the two cases is demonstrated to be very different, thus reinforcing the negative answer to RQ4. Fourth, the only small group of vocabulary (Acc. direction feminine new) consisting of 32 items (from the total of 1132 case forms for Test 2), where all confounding variables produce a negative effect, cannot be considered acquired by SLA standards and only equates to a third degree by HE standards, thus further reinforcing the importance of considering the additional factors of gender, case contexts and familiarity of lexis in case inflection production.

6.10. Results for the RQ6 - Working memory statistics and its correlations

6.10.1. WM correlations with overall success rates for the three tests

To answer RQ6, examining the role of WM in the success of inflection production, participants' WM span was measured (see 5.4.6) and correlations of those measurements with participants' success rates were analysed.

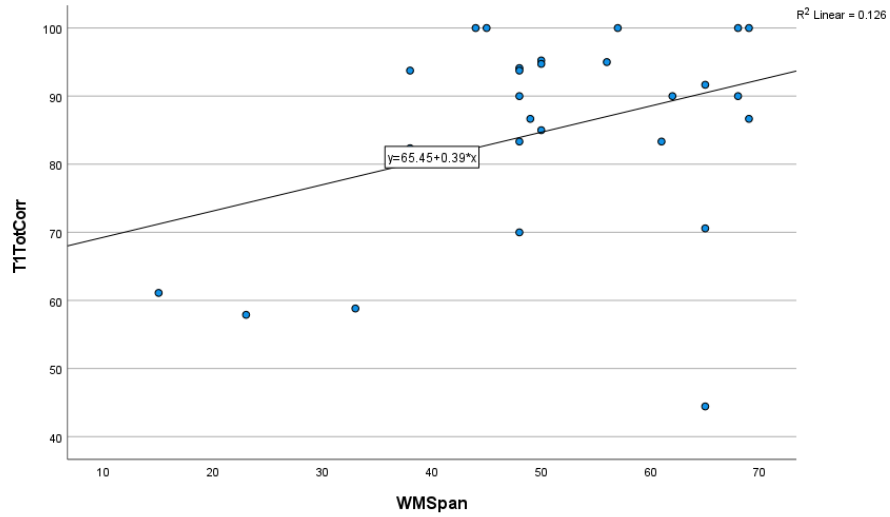
The average OSPAN Absolute Score for the participants was 51, with minimum span of 15 (P2019) and maximum of 69 (P2012 and P2027). However, there were three participants whose WM span was quite low (33 for P2015, 23 for P2017 and 15 for P2019). On the other hand, eight participants had WM span considerably above average (over 60.) These will be discussed in more detail in 7.4.2.

Pearson's correlation test was run, examining the correlation between the overall percentage of correctly inflected case forms in Test 1 and participants' WM. Pearson's correlation coefficient of 0.355 demonstrates that there is some moderate correlation between the size of WM and the success of inflection production in Test 1, but it is at the lower end of the scale. The p-value of 0.035 indicates that the correlation is significant.

The scatter plot in Figure 34 illustrates the above correlation with the fit line demonstrating moderate relationship between the case inflection production success and WM.

Figure 34

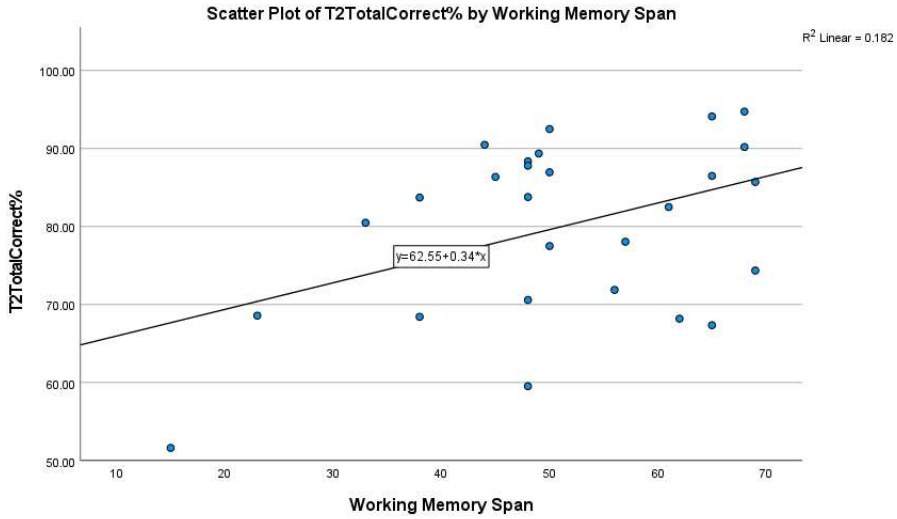
Scatter Plot Illustrating Correlation Between WM and Overall Success Rates for Test 1



For Test 2, Pearson's r-coefficient is 0.426 was slightly higher but also indicated a correlation with overall success rates which moderate. With significance value of 0.027, this is also a significant relationship. The scatter plot in Figure 35 gives the visual illustration of the positive linear correlation between the variables.

Figure 35.

Scatterplot Demonstrating Correlation Between WM and Overall Success Rates for Test 2



Unlike in the first two tests, overall success rates for Test 3 had a very weak correlation with participants' WM capacity. Pearson's correlation coefficient was calculated at 0.126; with p-value of 0.668, this correlation was not significant either.

From the above, it can be stated that no strong relationship was found between WM and the overall case inflection success rates. However, the analysis in this section demonstrates that some moderate and weak relationships with WM are observed at different stages of case inflection learning (with the highest *r* for Test 2), but they appear to be quite inconsistent. Therefore it can be suggested that some relationship between can exist but no straightforward answer to RQ6 can be offered at this stage.

6.10.2. WM correlations with gender, as a confounding variable

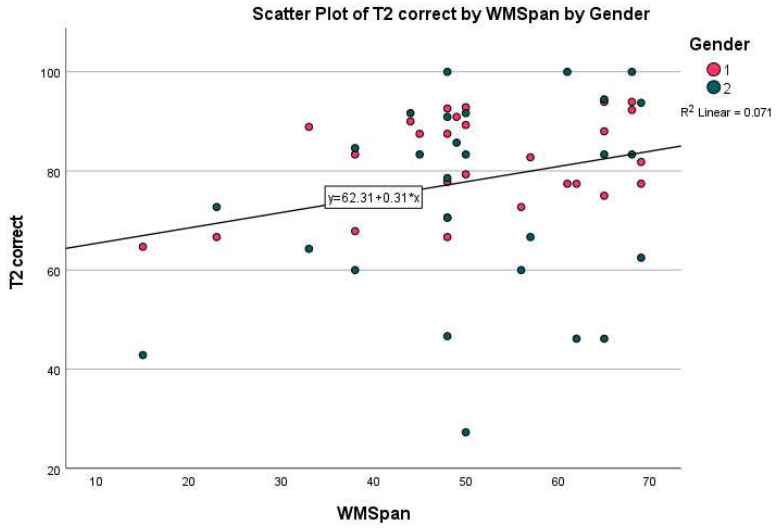
Furthermore, as gender was identified as one of the factors affecting inflection production, the correlation of success rates with WM were calculated for each gender set in each of the three tests.

For Test 1, the Pearson's Correlation between WM and gender was quite weak for feminine (r = 0.287) and almost non-existent for masculine (r = 0.014); both were insignificant, with p-values of 0.146 for feminine and 0.946 for masculine.

When gender was brought into Test 2 correlation (see Figure 36), it can be clearly seen that there are more feminine dots at the bottom of the graph.

Figure 36

Scatter Plot Demonstrating Distribution of Masculine (1) and Feminine (2) Correct Scores

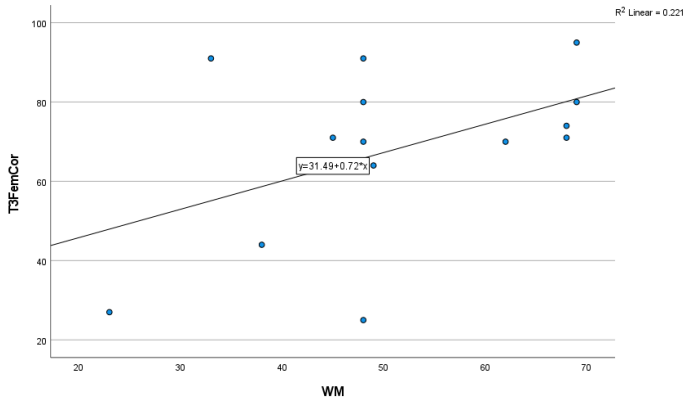


Running the Bivariate Correlation has produced the result opposite to the expectations. It was found that the correlation was only significant for masculine (r=0.393 and p=0.043); and quite weak (r=0.243) and not significant (p=0.222) for feminine.

In Test 3, opposite to Test 2, the correlation of WM with masculine scores was insignificant, as well as negative (r = -0.229), while for feminine - moderate, though still not significant (r= 0.470). The scatter plot in Figure 37 produced a steeper fit line for feminine scores, than that for the overall scores for Test 3, demonstrating a marginally stronger relationship.

Figure 37

Scatterplot Illustrating Correlation Between WM and Correct Scores for Feminine in Test 3



Similarly to 6.10.1, no strong correlations were found in this part of the analysis. The relationships between WM span and gender tend to be insignificant, except that with masculine in Test 2. Also, correlations appear to be quite inconsistent and range from almost non-existent to upper end of moderate, with the strongest correlation demonstrated for feminine in Test 3. Overall, the results are deemed inconclusive and no definite answer to RQ 6 can be given.

6.10.3. WM correlations with case inflection success rates for familiar and new vocabulary

With regard to familiarity of lexis, it was expected that the most challenging task of inflecting new vocabulary for a particular case would depend directly on the size of the WM, as it would require more cognitive resources (see 3.1). That is why, the correlation between WM and the success rates for new and familiar vocabulary was examined. Again, contrary to the expectation, some moderate correlation was found for familiar vocabulary ($r=0.392$ and $p=0.043$), which does not offer a conclusive answer to RQ6.

Though some correlations between WM and other variables were found, the results of for the three tests appear to be inconclusive, as no strong correlations were found and many were not significant, while moderate or weak correlations do not appear to be consistent and vary from test to test. Therefore it appears that no definite answer could be provided for RQ6. This will be discussed in 7.4.1.

6.11. Results for RQ6 - individual variability across the three tests and its relationship with WM.

As the analysis of WM correlations with various variables in 6.10 did not bear conclusive results, in order to further investigate RQ6, in this section, I examine individual success rates from the point of view of participants' WM characteristics. I will look at how these changed across the three tests that were conducted in the present study, for each case, for each gender, as well as for each individual case suffix under investigation. In addition, I will look at both group scores and individual trajectories. For the consistency of my comparisons, I only discuss the success rates of the 14 participants who completed all three tests.

6.11.1. WM and individual overall success rates

To start with, I am presenting the summary of the total success rates for the 14 participants for each test – see Table 37. For convenience purposes, the participants are listed in the order of their WM span – from highest to lowest. Also, the data for three participants with the lowest WM spans (P2001, P2015 and P2017) are highlighted in pink and the scores of four learners with the highest WM spans (P2008, P2012, P2027 and P2029) are in green.

Table 37

Individual Success Rates Across the Three Tests (T1, T2 and T3)

Participants	WM Span	T1 overall	T2 overall	T1T2 overall difference	T3 overall	T2T3 overall difference	Overall Difference
P2012	69	100	74	-26	88	+14	-12
P2027	69	87	86	-1	95	+9	+8
P2008	68	90	95	+5	74	-21	-16
P2029	68	100	90	-10	72	-18	-28
P2028	62	90	68	-22	64	-4	-26
P2002	49	87	89	+2	74	-15	-13
P2007	48	94	88	-6	68	-20	-26
P2018	48	83	60	-23	67	+7	-16
P2021	48	70	71	+1	76	+5	+6
P2023	48	94	84	-10	95	+11	+1
P2013	45	100	86	-14	78	-8	-26
P2001	38	94	68	-26	65	-3	-29
P2015	33	59	80	+21	90	+10	+31
P2017	23	58	69	+11	73	+4	+15

As, in Test 1, WM was found to moderately correlate with participants' success rates (see 6.10.1.), it is logical that two participants with the smallest WM capacity (P2015 with WM span of 33 and P2017 with WM span of 23) produced the lowest scores for Test 1 (59% and 58%, respectively). However, the next WM span of 38 (P2001) yielded 94% accuracy in Test 1. As it can be seen from Table 37, the variability at the top end of Test 1 is even higher and the highest success rates do not necessarily correspond to the largest WM span. For example, 100% accuracy was achieved by two of the participants highlighted in green at the top of the list (P2012 and P2029), as well as by P2013 with the average WM span of 45. On the other hand, one of the top WMs (P2027) did not make it to 90%, which was easily reached by two average WMs (P2007 and P2023). The only conclusion that appears to be possible to make for Test 1 is that the lowest inflection scores seem to be produced by the

lowest WMs, as well as the average success rates tend to correspond to average WM spans. This is not necessarily true the other way round. Also, a similar connection does not seem to exist for the highest scores.

Furthermore, as it was established in 6.10.1, the correlations between WM and overall scores got slightly stronger in Test 2 (but still moderate), while, in Test 3, the relationship got weaker. The individual rates do not seem to clearly reflect these changes. It appears particularly interesting in Test 3, which was a delayed test after participants were not taught for 6-8 weeks and some decrease in rates is normally expected. Table 37 shows that it is only true for half of the participants (7 from 14), while the other half produced an increase.

The size of both gain and drop between the tests varies a great deal and does not appear to necessarily correlate with WM. For example, the largest drop between Test 1 and Test 2 (-26) occurred for one of the bottom three WMs (P2001), as well as for one of the top four WMs (P2012). However, the largest gains in Test 2 were still produced by participants with the smallest WM span (P2015 and P2017), +21 and +11 respectively. These two participants carried on gaining in Test 3, though their gains were not the largest any more. At the same time, the success rate for P2029, who is one of the top Test 1 scorers (as well as one of the highest WM spans), deteriorated drastically, with -10 for Test 2 and -18 for Test 3. The other two highest WMs behaved in opposite ways – P2008 gained +5 in Test 2 and dropped -21 in Test 3 (the largest drop in Test 3), while P2012 first dropped -26 (one of the two largest drops in Test 2) and then gained +14 in Test 3. These three patterns of the highest WMs (“drop-drop”, “gain-drop” and “drop-gain”) can be observed in the performance of those with average WM spans too, though the figures vary. The most consistent accuracy was produced by P2021 (+1; +5), who has an average WM span (48) but was the only other participant (apart from the two lowest WMs) displaying the “gain-gain” pattern. Contrary to the expectations, learners with the lowest WM span appear to be consistent improvers and demonstrate “gain-gain” pattern in every subsequent test, producing the largest gains among participants.

6.11.2. WM and individual success rates for each case, gender and each case suffix

To contribute to RQ4, as well as to RQ6, the dynamics of individual success rates within each case set was analysed and revealed some more interesting tendencies. As it can be seen from Table 38, for the majority of participants, the changes between the post-test (Test 2) and the delayed test (Test 3) for both cases, appear to develop in the same direction –

either “gain-gain” or “drop-drop” (with the exception of only two participants – P2013 and P2001). (Please, note that there was no interim test for Acc.).

Table 38

Individual Success Rates for Each Case Across the Three Tests (T1, T2 and T3).

Participants	WM span	Prep.						Acc.		
		T1	T2	T1T2 difference	T3	T2T3 difference	total difference	T2	T3	total difference
P2012	69	100	75	-25	83	+8	-17	73	94	+21
P2027	69	87	83	-4	90	+7	+3	89	100	+11
P2008	68	90	93	+3	69	-24	-21	100	81	-19
P2029	68	100	97	-3	71	-26	-29	78	74	-4
P2028	62	90	67	-23	65	-2	-25	71	63	-8
P2002	49	87	90	+3	70	-20	-17	89	80	-9
P2007	48	94	81	-13	48	-33	-46	100	94	-6
P2018	48	83	54	-29	63	+9	-20	69	71	+2
P2021	48	70	73	+3	75	+2	+5	67	78	+11
P2023	48	94	75	-19	91	+16	-3	100	100	0
P2013	45	100	88	-12	93	+5	-7	85	53	-32
P2001	38	94	72	-22	65	-7	-29	62	65	+3
P2015	33	59	88	+29	92	+4	+33	69	88	+19
P2017	23	58	90	+32	95	+5	+37	36	47	+11

Thus, after the instruction period is completed, if there is a gain in scores for Prep., there is normally a gain in Acc. and vice-a-versa (see Table 38), though the amount of gain or drop can be different for the two cases. What is interesting is that a consistent gain is demonstrated by the two lowest WMs, as well as by the two highest WMs (with considerably larger gains for Acc). The other two highest WMs (P2008 and P2023) considerably dropped their scores for both cases. This piece of analysis points us towards the conclusion that, in delayed testing, individual success rates for each case tend to develop in one direction – they either both increase or both decrease.

As gender was found a crucial factor in case inflection production (see 6.7), gender effects were investigated within individual success rates (see Table 39 below). On one hand, the data supports the conclusion made in the section above regarding consistent improvement of participants with the lowest WMs. For masculine, the lowest WMs kept improving – P2015 produced the only gain (for this set) between Test 1 and Test 2 (+27), as well as the largest

overall gain, while P2017's main gain in masculine occurred between Test 2 and Test 3, but is comparable to that by P2015 (+23). P2023 is the only other participant who increased their overall scores within the masculine data set, with the other 11 dropping their scores to a different degree.

Table 39

Individual Case Inflection Success Rates for Masculine and Feminine Across the Three Tests (T1, T2 and T3).

		Masculine						Feminine					
Participants	WM span	T1	T2	T1T2 differ	T3	T2T3 differ	Total differ	T1	T2	T1T2 differ	T3	T2T3 differ	Total differ
P2012	69	100	77	-23	90	+13	-10	100	63	-37	80	+17	-20
P2027	69	100	82	-18	95	+13	-5	80	94	+14	95	+1	+15
P2008	68	100	92	-8	75	-17	-25	75	100	+25	71	-29	-4
P2029	68	100	94	-6	71	-23	-29	100	83	-17	74	-9	-26
P2028	62	88	77	-11	63	-15	-25	100	46	-54	70	+24	-30
P2002	49	100	91	-9	79	-12	-21	83	86	+3	64	-22	-19
P2007	48	100	88	-12	68	-20	-32	100	91	-9	70	-21	-30
P2018	48	100	67	-33	80	+13	-20	67	47	-20	25	-22	-42
P2021	48	100	71	-29	75	+4	-25	0	71	+71	80	+9	+80
P2023	48	86	78	-8	97	+19	+11	100	100	0	91	-9	-9
P2013	45	100	88	-12	83	-5	-17	100	83	-17	71	-12	-29
P2001	38	100	68	-32	79	+11	-21	86	60	-26	44	-16	-42
P2015	33	63	89	+27	90	+1	+28	56	64	+8	91	+27	+35
P2017	23	71	67	-4	90	+23	+19	60	73	+13	27	-46	-33

On the other hand, within the feminine set, after considerably smaller gains between the first two tests, the lowest WMs scores between Test 2 and Test 3 developed in opposite directions – P2015 made a huge leap from 64% to 91%, with +27 positive difference, while P2017 went down from 73% to 27% with a negative difference of -46.

However, this time, neither of the two produced the largest change, as P2021 (with an average WM) excelled from no correct feminine case inflection in Test 1 (0%) to 80% success rate in Test 3, which is the biggest progress in the entire experiment. At the same time, two participants (P2001 with low WM and P2018 with average WM) produced the largest overall negative difference of -42.

To summarise the above, there is one distinctive difference between genders with regard to case inflection – case success rates for masculine tend to decrease after Test 1 (except the two participants at the bottom of the list), which supports the conclusion regarding gender given in 6.7. The second tendency regarding participants with the lowest WM, who tend to have lower rates at the start but normally keep improving (except the final drop in success rates for P2017 for feminine in Test 3), is in line with the results in 6.11.1.

6.11.3. Individual acquisition trajectories for each suffix

Finally, I examined individual variability in accuracy for each individual suffix (RQ5 and RQ6). The results are shown in Table 40. This appears to be the level of analysis where some logic in the inflection accuracy changes could possibly be suggested.

Table 40

100% Accuracy Rates for Each of Case and for Each Gender

Participants	WM span	Prep. Masculine T1	Prep. Masculine T2	Prep. Masculine T3	Feminine T1	Feminine T2	Feminine T3	Acc. masculine T2	Acc. masculine T3	feminine T2	feminine T3
P2012	69	100	73	84	100	75	75	82	88	50	83
P2027	69	100	76	92	80	100	86	92	70	86	100
P2008	68	100	89	63	75	100	86	100	100	100	57
P2029	68	100	100	65	100	90	88	80	75	75	64
P2028	62	88	75	62	100	43	100	82	89	50	63
P2002	49	100	90	67	83	88	80	92	100	83	56
P2007	48	100	81	47	100	83	50	100	100	100	83
P2018	48	100	58	69	67	43	33	88	100	50	20
P2021	48	100	71	70	0	78	100	69	85	63	70
P2023	48	86	67	94	100	100	75	100	100	100	100
P2013	45	100	83	89	100	100	100	93	64	67	29
P2001	38	100	68	81	86	67	29	67	75	50	56
P2015	33	63	100	91	56	67	100	73	100	60	86
P2017	23	71	94	100	60	80	75	13	100	67	0

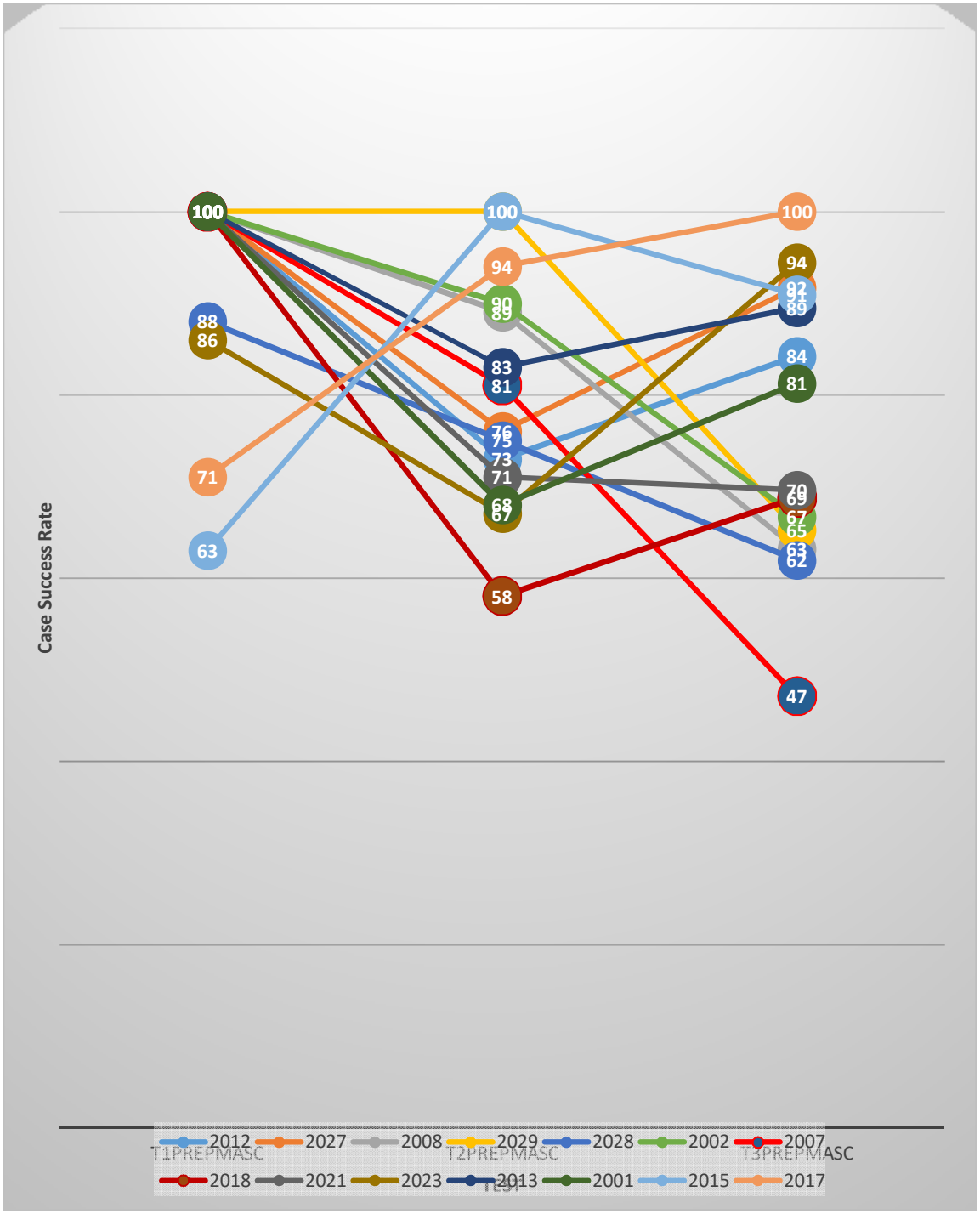
To start with, the 100% accuracy rates were marked and traced across individual samples. This allowed me to observe how learners develop accuracy for each of the inflections.

First, a large number of 100% correct inflection scores is noticeable for Prep. masculine in Test 1 – 10 from the total of 14 participants. From the remaining four, two participants were P2015 and P2017, who had the lowest WM spans and produced the lowest scores for Test 1 (Prep. only) for both masculine and feminine. However, if these two are followed across the three tests, P2015 reaches 100% accuracy in Prep. masculine by the time of the post-test (after the introduction of Acc.), while P2017 produces the same level of accuracy for this inflection in the delayed test (Test 3). This means that they still achieved the desired accuracy but with the delay, which could be logically explained by their lower WM capacity. The introduction of Acc. did not seem to affect their progress in Prep, where their cognitive resources, however limited by WM, were directed. The remaining two participants (P2023 with an average WM and P2028 with a fairly high WM span) did not reach 100% level for Prep. masculine within this experiment, though P2023 got quite close, producing 94% in Test 3. Individual differences will be discussed in 7.4.3.

One more interesting observation is that the top rates in Prep. masculine are rarely retained. This is in line with the previous report on the drop of Prep. accuracy (see results in 6.4.2. and discussion in 7.2.1), but a fair amount of variability can still be observed here. The line graph in Figure 38 gives a visual representation of how the Prep. masculine success rates changed within individual samples throughout the three tests. It is clear that the drop in accuracy for Prep. caused by the introduction of Acc. in Test 2, is most common tendency (including nine of the top ten scorers). The exceptions are the two lowest WMs (P2015 and P2017) who had a low start but improved significantly and consistently across the three tests, and P2029 who managed to retain the top score in Test 2 before drastically dropping it in Test 3. The two largest drops though, are the two participants with the average WM spans – P2007 and P2018, whose individual characteristics will be discussed in 7.4.3.

Figure 38

Dynamics of Individual Rates for Prep. Masculine Inflection for the Three Tests (participants are listed in the legend from the highest WM span to the lowest)



Moving on to Prep. feminine now, the two participants who never achieved a 100% accuracy for Prep. masculine (P2023 and P2028), delivered this level for Prep feminine in Test 1. Furthermore, four other testees produced 100% scores for Prep. feminine in the first test – two with the high WMs (P2012 and P2029) and two with the average WMs (P2007 and P2013). Moreover, P2008 and P2027 (both high WMs) produced 100% correct inflection in Prep feminine in Test 2, while P2015 and P2021 (low and average WM respectively) reached it in Test 3. Altogether, ten participants had a 100% success rate for Prep. feminine,

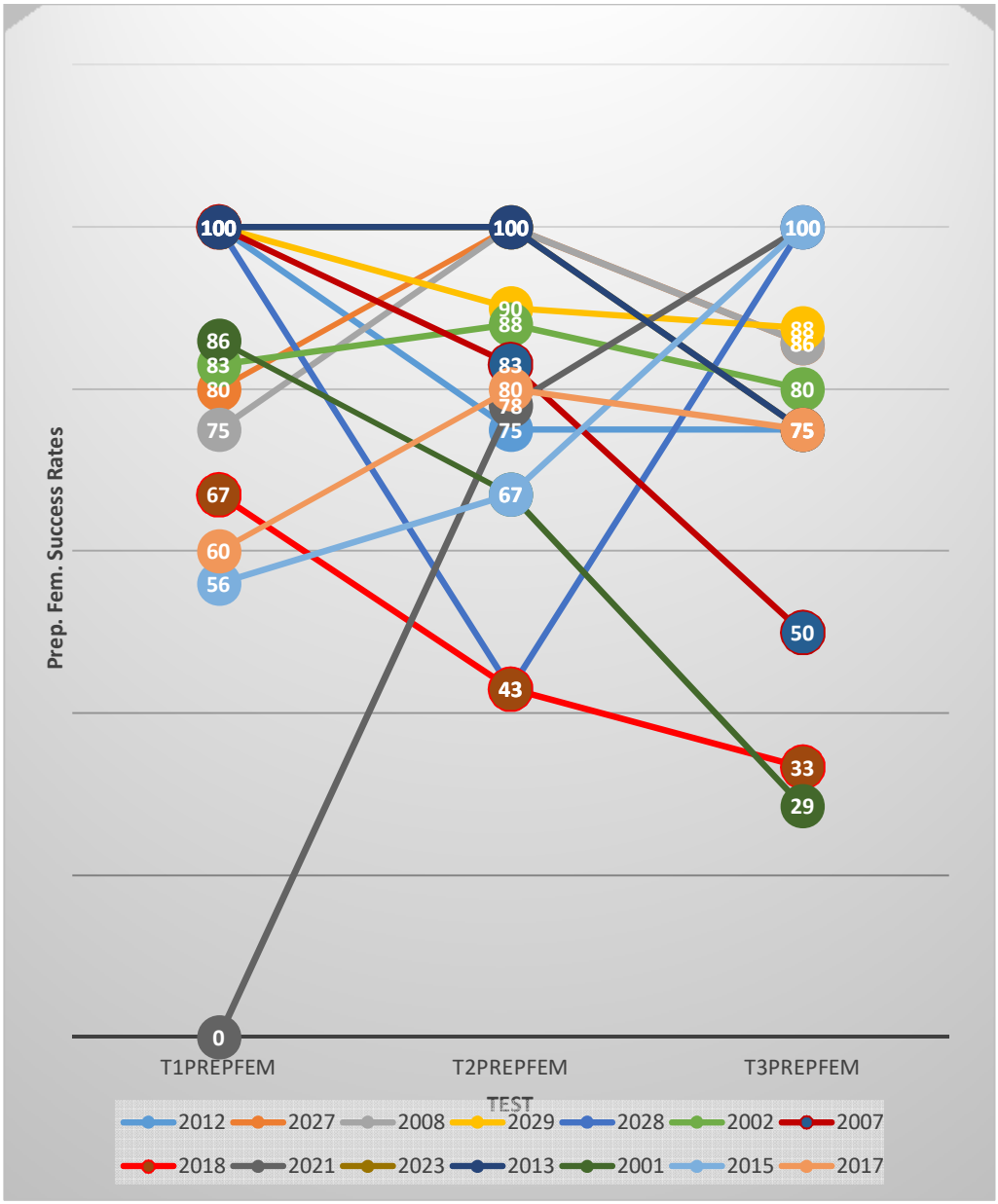
which is the same number as for Prep. masculine, but more of them achieved it with a delay in time. This could be due to the two-step inflection production of feminine case forms, which is assumed to require more cognitive resources during learners' processing, and therefore could take longer to master (see discussion in 4.1).

With regard to retainment, two testees (P2013 and P2023) managed to maintain the 100% correct production for Prep. feminine in Test 2, with only the former keeping it in Test 3, which is more successful than for Prep. masculine. Though it was established in 6.7 that the case inflection behaves differently for different genders, the line graph displaying individual trajectories (see Figure 39), is able to demonstrate visually why the spread of data was so wide in feminine (see Figure 20). Though a few lines go down indicating the accuracy drop (five participants with more than -20 decrease in accuracy – P2012, P2007, P2018, P2023 and P2001), there are two steep climbs for P2021 with 100% increase and a huge increase of 44% for P2015 (one of the lowest WM). The introduction of Acc. after Test 1 does not appear to affect the accuracy for Prep. feminine as much, as it impacted the masculine scores (except P2028 who displayed a sharp "drop-gain" pattern).

Figure 39

Dynamics of Individual Rates for Prep. feminine Inflection for the Three Tests

(listed in the legend from the highest WM span to the lowest)



The above results demonstrate that the absolute majority of participants manage to reach 100% accuracy for Prep. but at different stages of their learning, that is, in Test 1, Test 2 or Test 3.

To some degree, similar phenomenon occurs with accuracy for Acc. masculine, but there are also some noticeable differences (see Table 40). For example, half of the participants (7 from the total of 14) produced the 100% accuracy (which is only three less than for Prep. masculine), but only three of them manage it in Test 2, the first test since the introduction of Acc. The other four testees reached this level only in the delayed test, while half of the participants never achieved it in this experiment. These developments are discussed in

7.2.1. It has to be noted that Acc. was taught only for five weeks, unlike Prep., which was in Production for ten weeks.

The most insightful observation though is that there appears to be a slight re-distribution of higher and lower scores between the four suffixes within individual samples, when the rates increase for one suffix and decrease for the other in the same test. To illustrate this, I have created line graphs showing the changes in the scores for five participants - one with a high WM span (P2027), two participants with a low WM spans (P2015 and P2017), and two with an average WM capacity (P2002 and P2017). From these, P2027 and P2002 were high scorers, P2017 was one of the low scorers, while the accuracy of P2013 and P2015 has undergone large changes, going in opposite directions, throughout the three tests.

Figure 40
Dynamics of Success Rates for the Four Inflections for P2027

Figure 41
Dynamics of Success Rates for the Four Inflections for P2002

Figure 42
Dynamics of Success Rates for the Four Inflections for P2013

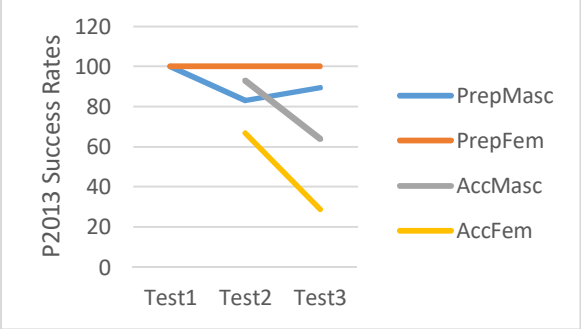


Figure 43
Dynamics of Success Rates for the Four Inflections for P2017

Figure 44

Dynamics of Success Rates for the Four Inflections for P2015

The figures above, together with Table 40, suggest four main tendencies in case inflection acquisition regarding gender, which will be discussed below.

1) Each case suffix had its own acquisition trajectory (see Table 16). This suggests that masculine and feminine suffixes for the same case can have different, and sometimes opposite (see Figures 34, 35 and 37), acquisition trajectories, even when the suffix is the same for both genders like “e” in Prep.

2) Consequently, at one time, masculine and feminine suffixes for the same case can have different, or even very different, levels of accuracy. In addition to the results in 6.7.2, showing that there can be a significant difference in case accuracy between genders in group scores, the analysis above demonstrates differences at the level of individual scores. Two extreme examples (see Table 40) are P2021, who, produced 100% accuracy for Prep. masculine and 0% correct scores for Prep. feminine, in Test 1; and P2017, who supplied 100% for Acc. masculine and 0% for Acc. feminine, in Test 3, (as well as 100% Prep. masculine – to confirm the distinction between the two masculine suffixes). Among less drastic differences, P2001, whose success rate for Prep. masculine in Test 3 was 81%, produced only 29% correct for Prep. feminine, in the same test. These differences are not necessarily restricted to lower WMs – P2012 (one of the two highest WMs and one of the highest scorers in the experiment), in Test 2, supplies 82% for Acc. masculine and only 50% for Acc. feminine. Similarly, P2008 (WM span of 68) delivered 100% accuracy for Acc. masculine in Test 3, with only 57% for Acc. feminine. Four exceptions from this tendency are found in Test 1 (where Prep. was the only case in production), when P2012, P2029, P2007, P2013 produced 100% success rates for both Prep. masculine and Prep. feminine (see Table 40). Also, three participants (P2008, P2007

and P2023) supplied 100% rates for both genders for Acc. in the Test 2 (the first test on Acc.). These are further analysed in the two paragraphs below.

3) *At the offset of case learning*, masculine accuracy tends to be higher (see Figures 34, 35, 37 and 38, as well as Table 40). Figure 42 shows suffix trajectories for one of those few samples listed above (P2013), where both masculine and feminine case forms were initially produced simultaneously at 100% rate. Only three samples in Table 40 demonstrate the reversing of case accuracy for genders at the start of learning (when the initial feminine case inflection success rate is higher than the masculine for the same case) - P2028 and P2023 for Prep.; and P2017 for Acc.

4) Within the analysis of the distribution of 100% scores in Table 40, one more tendency is arising. If a participant manages to retain one case suffix at a high level of accuracy, then the scores for other suffixes tend to go down, as it is clearly illustrated in Figure 40, Figure 42 and 37. Often, at the next stage of learning, the initial high rate for that particular suffix drops, while the accuracy for another suffix (or other suffixes) increases. Also, Figure 41, demonstrates a slightly different re-distribution of success rates - P2002, in Test 2, managed to produce all for inflections at a similar level at a expense of the drop in Prep. masculine success, with the drastic decrease in accuracy for Prep. masculine and Acc. fem in Test 3. Similarly, 100% accuracy for both genders in the same case (see results for top rates for Test 1 in point 2 above) at a cost of either following reduction in success rates for Prep. masculine (P2013), Prep. feminine (P2029), or both (P2012 and P2007).
   These results will be discussed further in Chapter 7.4.

5) Finally, the qualitative analysis of the individual samples (rather than the scores), employing the pairing up of Prep. and Acc. case contexts, built into Comics Tests' design, suggests that, at the offset of case learning, masculine forms for a particular case tend to emerge correct in learner's production first, that is before feminine for the same case. Example 1 in 6.7.1, as well as Example 5 in 7.3.3, also clearly demonstrate this phenomenon. (This is also in line with the point 3 above.) That is why the presence of correct Acc. feminine suffix, which is overt, can assist in marking of Acc. masculine zero-inflection (which is syncretic with masculine base form) as correct. However, the absence of correct feminine inflection in Acc. cannot be used as the indication of a base-error for masculine, as masculine case forms tend to be produced correct first.

.

6.12. Results for RQ7 - qualitative data analysis

The consequential research design of the present study included three qualitative instruments, namely, Learners' Diaries, Qualitative Questionnaire and In-depth Interviews (see 5.5.1, 5.5.2 and 5.5.3). These were designed to enable me to answer RQ7 (see 5.1.1), which investigated learners' perceptions of spiralling instructions, Russian case inflection and any learning strategies, specific to inflection, should they have any, as well as possibly contributing to exploring learner variability (RQ6).

6.12.1. Qualitative data collection and coding

The four Learner Diaries varied in length considerably. For example, one of the volunteers (P2001) stopped after three weeks and then wrote a bit more towards the end of the Instruction period, providing very limited data. Two participants (P2004 and P2027), on the contrary, were fairly consistent in filling their notes and provided eight and twelve A4 pages of material, though weekly entries differed from one line to half of A4, and some included pictures. The remaining participant (P2029) provided 3.5 A4 pages (after they were typed up). Altogether, the four learner diaries supplied data that added some value to the other data collected, as well as helping me design the other two qualitative instruments (see 7.5).

With regard to the Qualitative Questionnaire, 24 participants (out of 27 who completed the instruction period) filled in the online Part 1, and 21 of them have returned Part 2. This is an excellent response, considering that participants had already finished learning Russian with me by the time they were asked to fill this in. Some participants were quite generous with their comments, while others filled in only the compulsory sections. The results of Part 1 (anonymous) were summarized with the help of Jisc website, while Part 2 had to be summarized manually.

The six In-depth Interviews varied in length between 40 and 50 minutes, depending on how much detail the interviewee was willing to go into. They were recorded on my laptop, then transcribed and coded.

The coding was done manually, but the approach was slightly different for Learner Diaries, for the Qualitative Questionnaire responses and for the In-Depth Interviews. First, as the Learner Diaries were unstructured, I examined the instances when the case suffixes were

mentioned, in order to identify 'noticing'. Some common themes which were relevant to RQ7 were identified and the words which belonged to those themes, were marked.

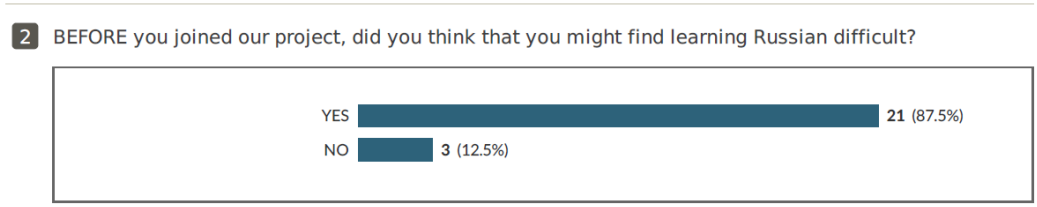
During the coding process, it was decided to merge the "vocab" and "words" with "pronunciation" and "Cyrillic" items into one code, as these two aspects were not investigated in the present study. The only reason why the Learner Diaries were coded for these, is to compare the number of instances when difficulties of learning Russian were associated with inflection as opposed to vocabulary (or "words") retrieval or production. Due to the limited number of relevant items found in Learner Diaries, the colour coding was chosen. The coding protocol for Learner Diaries was as follows:

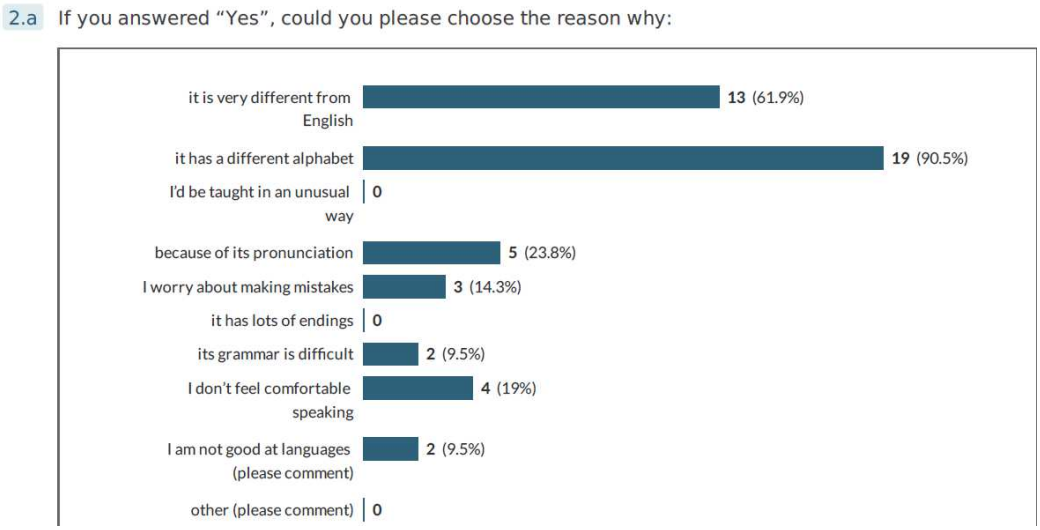
	Confident
	Fun, positive, enjoy
	Hard, difficult
	Suffix, grammar, case
	Vocab, words, Cyrillic, pronunciation

Second, Part 1 of the Qualitative Questionnaire (see 5.5.5), which was anonymous, was summarized through the Jisc website, which provided percentages of respondents who selected a particular choice for each of the questions, as well as the answers for the open-ended questions (see examples in Figure 45). The most relevant percentages were then transferred into tables.

Figure 45

Extract from the Jisc Summary for the Qualitative Questionnaire (Part 1).





Par 2 of the Qualitative Questionnaire, which was not anonymous (see 5.5.5), was coded separately. Initially, I tried to match the codes to the questions, for example, Question 3 was asking about building an L2 system; thus, the code of A3 was assigned to any mention of a linguistic system, that is, "system of the Russian language", "Russian grammar system", "system of suffixes", "systems of grammar forms" and similar. However, during the coding process, I realised that codes need to reflect some more detailed nuances of participants' perceptions. For example, Question 7 probed how learners perceive the differences between nominal and verbal inflection. It has become apparent that the difference needs to be made between statements "I find noun suffixes more difficult than verb suffixes" and "I think of noun suffixes more than verb suffixes", as some participants answered "yes" to the first, but then commented that they concentrate on the verb conjugation more (e.g., P2001, P2026). Consequently, though the codes are related to the questions in Part 2 of the Questionnaire, they do not match the question numbering, as there are only twelve questions, but 22 codes. The list of codes can be found in Appendix J.

Finally, the semi-structured format of the In-depth Interviews, which were based on the interviewees' answers to the Qualitative Questionnaire, as well as Learner Diaries of some of them, allowed me to use the Questionnaire Code as a basis for the Interview coding. One of the participants who handed in their Learner Diary, could not be interviewed due to them starting a new job. Two interviewees did not volunteer to write a Learner Diary, but were chosen for an Interview, as one of them had the lowest WM span, while the other was one of the top four WMs, as well as one of the top scorers for Test 2. The six In-depth Interviews were transcribed with the help of speech-to-text software (see 6.1.2) and analysed individually.

6.12.2. Results of the analysis of Learner Diaries

Despite the inconsistency and different amount of notes, Learner Diaries provided enough material to observe certain tendencies. The results of the analysis of the Learner Diaries material are presented in Table 41.

Table 41

Results of the Coding for Learner Diaries

	Confident	Positive (fun, enjoy)	Hard (difficult)	Suffix (grammar)	Vocab, pronunciation
P2001	0	2	1	2	15
P2004	8	7	5	1	9
P2027	1	3	18	9	6
P2029	0	12	19	5	6

Three main themes emerged in the analysis of these data. First, all volunteers commented on the positive perception of the instruction that they received, though some provide more of these comments than others – between 2 and 12. Second, all diaries contain references to difficulties in learning Russian, with the amount of commentary, again, varying between 1 and 19 instances. Third, there are considerably more mentions of vocabulary and pronunciation (between 6 and 15), than those of suffixes and grammar (between 1 and 9). Finally, it has become apparent that the confidence theme was only important for one participant (2004) and, therefore, was disregarded in further analysis. These results are discussed in 7.5.

6.12.3. Results of the analysis of Qualitative Questionnaire

The results for Questions 1-7 of Part 1 (anonymous) of the Qualitative Questionnaire are summarized in Table 42 (see the detailed questions in Appendix G). Questions 8, 9 and 10 were open-ended and mainly contained thanks to the instructor, positive comments with regard to participants' learning experience and high evaluation of the level of teaching. The only two suggestions were to provide worksheets online (which was done in the last four weeks) and to increase the length of the sessions or the course (which was not logistically possible, as discussed in 5.3.1).

Table 42

Summary of Results for the Part 1 of the Qualitative Questionnaire (Anonymous)

Q1	Interest in Russia/Russian	General interest	Love learning new things	Like a challenge	Wanted to try a new language
Reasons for joining project	2	12	14	10	15

Q2	Yes	Different alphabet	Difficult pronunciation	It has lots of suffixes
Preconceived difficulty of learning Russian	21 (87.5%)	19 (90.5% of positive responses)	5	0

Q3	informative	helpful for my understanding	Crucial for my learning	Easy to follow	Unnecessarily repetitive
Instructor's explanations were	20	20	13	16	0

Q4	essential	helpful for my grammar	useful for understanding how Russian works	challenging	not necessary
Speaking activities	17	13	18	16	1

Q5	Not affected at all	Affected in the worst possible way	Affected mildly	Affected quite a lot	Affected severely
Effects of transferring online due to COVID	3	0	13	8	0

Q6	Gender agreement	Pronunciation	Cyrillic	Using noun suffixes	Understanding noun suffixes
Perception of difficulty *after* the course for the majority	Under 5	50:50	Under 5	Over 5	Under 5

Q7	Did not address	Not enough time given	appropriately	Too much time given
How the above aspects were addressed	0	1-5 (3 for suffixes; 5 for pronunciation)	18-24 (for different aspects)	1-2 (Cyrillic and pronunciation)

As is clear from Table 42, the absolute majority of participants did not have any specific interest in Russian or Russia (25 from the total of 27), and only just over a half were

interested in learning a new language (15). The majority joined the project because they liked learning new things and because of general interest (see Q1). This is in line with the criteria which was determined for the selection of the participants, who were not from language departments, were not heritage speakers and were not exposed to Russian in the past.

Not surprisingly, the majority of learners perceived Russian as a difficult language to learn (21) (see Q2). However, the difficulties were connected with the different alphabet (19) or pronunciation (5), while suffixes or grammar were never mentioned among the expected challenges (0). Interestingly, after completing the course, the absolute majority of participants rated learning Cyrillic as a considerably lower level of difficulty (under-5 rating) (see Q6). In contrast, the high levels of difficulty (5+) were selected by the majority of participants for using noun cases. At the same time, understanding noun suffixes was rated mainly at the under-5 level of difficulty. Pronunciation remained as one of the main challenges for half of the participants. The majority of participants (18-24) said that all the above aspects were addressed appropriately throughout the Instruction period (Q7). On one hand, five learners wished more time was spent on pronunciation and three would have welcomed more time for learning grammar. On the other hand, two participants thought that too much time was given to learning to read Cyrillic in the first half of the Instruction period (before cases were introduced) and only one was of the same opinion regarding pronunciation.

Furthermore, the response with regard to the quality of instruction that participants received is very positive (see Q3). Also, about a half or more (13 - 18) participants found speaking activities “essential”, “helpful for their grammar” and “useful for understanding how Russian works”, despite finding these “challenging” (16) (see Q4). Only one participant classed them as “not necessary”. One of the comments clearly stated: “speaking ultimately improved my learning much more effectively”.

Finally, Question 5 investigated participants’ perceptions of how their learning was affected by the transition to the online format with the start of the COVID-19 pandemic. None of the respondents thought that they were “affected severely” or “affected in the worst possible way”. Just about a half of the experimental group (13) considered the effect of the change “mild”, while eight participants (less than one third) said that they were “affected quite a lot” (see the exact descriptions of each category in Appendix G). Three participants stated that the transition did not affect them at all.

The results of Part 2 of the Qualitative Questionnaire are summarized in a table in Appendix K; the codes can be found in Appendix J and the full list of questions in Appendix H. The answers were often grouped and merged under certain codes and the most relevant results are presented below.

Though there was a lot of variation in responses, the analysis reveals certain tendencies. First, the majority of respondents report that they enjoyed learning languages at school, with only five stating that they did not. However, the five negative answers do not correlate with the participants' success rates in the present intervention or with their WM capacity, except perhaps for P2019, who has the lowest WM span in the group, as well as being the lowest scorer.

Second, almost half of the respondents feel that they have an emerging system of the L2 in their head (choice C), while the other half think that some parts of their knowledge is within a system, but some other parts have not found their place yet (choice B). Only one participant (P2015, one of the two lowest WMs) stated that they perceived their knowledge as being scattered (choice A). However, these choices did not consistently link to the success of case production.

Third, I believe that it would be true to say that the majority of the respondents demonstrate the qualities associated with successful learning, such as making comparisons (A4), looking for patterns (A5) and trying to understand the reasons behind things (A6), as well as connecting new information to something they already know (A21). For example, 16 respondents (from 21) stated that comparing Russian to other languages is "somewhat true" (3), "generally true" (4) or "always or almost always true" (5) of them, with only five selecting "never or almost never true" or " generally not true" of them for this question. There is a similar ratio for reasoning (17 : 4 respectively) and for linking new information to old information (19 : 2, respectively), while, for identifying patterns, only one participant chose a negative answer (P2001). This demonstrates that the majority of the respondents have similar learning strategies. However, it is noticeable that these strategies are not always applied to case morphology to the same extent, as, for example, the average rating for reasoning (A6) is higher than for looking for reasons for "why the suffixes are there" (A6a) – 3.6 : 3, respectively. In some cases, the difference is fairly large, e.g., P2017 (one of the lowest WM and of the lowest scorers) – 4 : 1, respectively. This shows that some participants think about reasons for inflection less, or considerably less, than generally when learning other items.

Fourth, despite the above, almost all respondents said that they notice case suffixes either "quite a lot" or "always try" to notice them (A7 and A10), both when reading and speaking, with the exception of P2019 again. Interestingly, all respondents stated that they noticed suffixes more in familiar words. Also, the results demonstrate that no participants feel that they find knowing the meaning of suffixes problematic (A8), though the degree of confidence differs slightly between individual learners (with rating from 1 to 3). In addition, A9 examined the level of importance that participants attach to meaning, as opposed to form. From 21 respondents, only four participants thought that learning words was more important than learning suffixes. At the other end of the scale, six marked that statement as "generally not true of me", thus acknowledging the prior importance of form. The remaining 11, which is just over a half tried to keep the middle ground. Thus, despite receiving the same instruction, which emphasized the importance of grammar form, the attitude to the importance of form varies.

Fifth, A11 aimed to tease out the differences in difficulty between verbal and nominal inflection. Over a half of the respondents (12) perceived case suffixes as more difficult than verb suffixes, with only four learners considering them easier to learn, while the remaining five thought these two were similar. Interestingly, only four participants said that they thought about noun declension more than about verb conjugation, while 11 stated the opposite. There appears to be a paradox when over a half of participants find case inflection difficult but at the same time many of participants do not think about it as much.

Sixth, learners' differences carry on manifesting themselves in prioritising different aspects determining the choice of inflection, when selecting an suffix (A12). For example, five respondents start with identifying the noun's gender, nine – with the verb that the noun is connected to, four think of the context first and two begin with the preposition. No pattern has been identified in these selections. Only seven respondents expressed an interest in developing a strategy for producing correct inflection.

6.12.4. Results of the analysis of In-Depth Interviews

To investigate the RQ& further, the six In-depth Interviews were coded and analysed. To start with, a brief summary of the interviewee data is presented in Table 43.

Table 43

Brief Summary of the Interviewee Data.

	P2001	P2004	P2008	P2019	P2027	P2029
Test 1	94	85	90	61	87	100
Test 2	68	87	95	52	86	90
Test 3	65	Not tested	74	Not tested	95	72
WM Span	38	50	68	15	69	68

The In-depth Interviews were designed on the basis of Part 2 (not anonymous) of the Qualitative Questionnaire and Learner Diaries (if applicable), thus provided further information for the codes analysed above (see Appendix L). During the course of analysis, in order to identify some distinct results which would inform the present study, the initial codes were merged into larger categories:

- conscious effort in building a linguistic system for L2 and the role of inflection in it;
- the view of how languages are learnt;
- directing attention to case inflection (noticing) and cognitive effort in understanding the inflection concept;
- designing strategies for mastering case inflection.

From the six Interviews, it has become apparent that the concept “the system of Russian” (or “language system”) is understood differently by different interviewees. For example, P2001 views language as a system of words, rather than that of grammatical forms, and, when describing the system of Russian that they have in their head, they explain how they organize vocabulary in different lexical groups and look for links and connections to remember words. A similar narrative emerges in the Interview with P2004, who states directly: “I think it’s all based around vocab and the specific kind of objects and words”. When asked directly about learning suffixes though, the two interviewees gave different answers - P2001 talked about verb suffixes, while P2004 concentrated on word order and sentence structure. Neither of them mentioned noun suffixes voluntarily during this part of the discussion. Eventually, P2001 revealed that they sometimes have suffixes “sit on top of” vocabulary groups in their head, e.g., “e” “sits” on top of “months”. In a similar way, P2004 said that they try to remember that some “vocab” has a particular suffix (“musical instruments” have “e”) and the other one does not (games); this still evolving around vocabulary, rather than a separate concept of case. Uncharacteristically, at the very end of their Interview, P2004 said that “grammar is important; you can’t really learn a language without grammar". P2019 (the lowest WM) is considerably less sure about what their system of Russian looks like, and the questions are frequently answered with “I don’t know.”, “I am not sure.” or “I did not think about it.” When probed further, they mainly discuss verb suffixes

and do not offer any kind of description of an inflection system they might have. On the other hand, P2008 said that they "quite like grammar", but their system was evolving around the sentence structure (similarly to P2004), rather than building a conceptual system of Russian cases.

P2027 (one of the highest WMs and one of the top scorers) was considerably more definite about their interest in "how languages are learnt" and "how different languages work", as well as writing in their Learner Diary that they "enjoyed trying to construct a system and comparing and contrasting it to other languages". Unlike other interviewees, P2027 is very clear on how they construct the system of Russian language in their head. Moreover, they talked about more than one system, for example, "a system of Russian sounds" and "a system of structures", in their words. By "structures", they mean short constructions, involving inflection, e.g., "v + *month* + e" (stands for "in + *month* + e*[Prep.]*"). They even keep a notebook with these constructions, linking them to each other by arrows, - in parts these connections appear similar to connections that I suggested in Figures 12a – 12d. They think of different elements of the linguistic system as pieces of jigsaw, though not one big jigsaw but rather several of those, which are connected to each other. Moreover, they say that they identify gaps in that jigsaw and look how to fill them in, saying that "it is nice to see it all click into place".

Finally, P2029, to start with, also appeared to have a fairly well-defined idea about their system of the Russian language. In their Learning Journal they wrote "I feel like I have a structure of Russian knowledge, with gaps to fill. And I am keen to fill them!". In the Interview, they develop this by describing their system as a stacking of coloured blocks. In reality, when probed further, they clearly referred to the sentence structure, rather than a system of grammar as a whole, and "the gaps to fill in" were, in fact, words that they didn't know; the stacking of the blocks seem to be quite linear. In addition to that, tenses were given as an example (though tenses were not studied within the Instruction period); this was illustrated with the word "will" as a potential gap-filler. Suffixes were not named as part of that system, but when asked about these directly, P2029 was not very sure about their place in the language system, did not have a particular colour or a shape assigned to them and said that they were "separate", as well as that they "struggle to keep them all in their brain". Then, they came up with the idea of suffixes possibly being "like a bolt-on" which is added to a word, and specified that they have to be a different colour from the word. Interestingly, the recurrent theme of "struggle", "worry", "concern" and "pressure" seems to be quite prominent in both the Interview and Learner Diary by P2029, despite their statement that they enjoy learning. These are normally linked to "difficulties" in mastering something. Difficulties are

also regularly mentioned by P2027; however, in contrast to P2029, they never mention "worry" or "concern", but tend to look for or offer some kind of a solution instead.

With regard to the actual learning process (the second category), P2001 and P2019 equate learning to memorization and prefer to speak using pre-learnt chunks, rather than trying to understand how Russian grammar system works and construct their own sentences. P2001 was practicing phrases from the handouts by saying them to members of their family (who do not speak Russian), while P2019 was making lists of vocabulary and learning those as a main learning strategy, despite being advised not to do that in class. P2004 and P2029 stress, on numerous occasions, that practice is key for their learning. For example, P2004 explained: "I would write out a list of different words and then …. Conversations, different sentences, structures, suffixes. For me, it is seeing it written down and then speaking it aloud… yeah, repetition of speaking it aloud". Though P2029 also said that they "repeat regularly what was done in class and re-read sentences", their learning sounds more conservative, as they "learn better in lists than spider diagrammes". P2008 appears to have a totally opposite idea about learning, as they tend to think that they "pick languages up naturally" and do not really practice in between sessions. They said that they produce their own sentences and "try not to think too hard, so I could say it with more fluency", thus attempting to appeal to their implicit knowledge, rather than relying on cognitive effort. Different again, P2027 sees learning as "understanding the system" and "having a go at using it". Two of their quotes illustrate these: "you need to know why a language does it [*changes a suffix*]"; and "I can try. And if I don't get it right – you would correct me.". This is in stark contrast with P2019's statement "It has to be perfect. If I am not sure that it is correct, I wouldn't say it.". In a way, this strive for perfection comes through P2029's data - in the Interview, they said directly that they "get very frustrated, when I don't get it right".

In relation to the third category (focusing on the noticing of inflection), the six interviewees, again, display some considerable differences. For example, though P2001 stated in their Questionnaire that they believe that suffixes are as important, as words (A9), their Interview suggests that their prime focus is on vocabulary, rather than on inflection, as all five examples that they gave (A3, A5, A6, A14 and A16) relate to learning words. Also, they thought that Spanish did not have any suffixes (while it does), which makes their claim that they "always try" to notice suffixes when reading and "always" think of them when speaking, quite speculative. In the Interview, they said that they tend to "think about suffixes more when they speak" than when they read, but then were absolutely clear about prioritizing verbal inflection over nominal. Moreover, P2001 confirmed that they "almost never" look for reasons behind linguistic aspects and said that they believed that these need to be

memorized. On the contrary, P2004 gave a clear explanation that they notice suffixes when they read but not when they listen to others, because when reading, they can go at their own pace, but that they are “so focused on listening to specific words that the suffixes just didn't really matter too much”. This appears to be true for P2029 too, but they also confessed that, even when they notice the suffixes, they do not necessarily think of why they are there. Again, unlike other interviewees, P2027 “looks out” for suffixes and if they don’t understand why it is there, they feel they need to find out. Participants answers in this category suggest that deliberate attention to inflection tends to contribute to higher scores, but sometimes participants claim that they notice suffixes when in reality they don’t.

In relation to the last category (designing strategies for producing inflection), P2001, P2004, P2019 and P2029 could not formulate any strategy for choosing a case suffix and were really struggling to come up with any kind of algorithm or a list of steps. For example, P2004 speculates: “I don’t think it was entirely subconscious. I definitely don’t think that. I think I was actively thinking … this word needs this … this suffix but I can’t remember exactly”. P2029 said that they sometimes try to apply new suffixes to different words, but gave an example of verb conjugation, rather than case declension, and then summed it up by saying that “it is all done by memory”. At the same time, they talk about “making up their own sentences” and “stacking” their “coloured blocks”, which seem mutually exclusive.
In addition, they said that they sometimes use mnemonic associations, for example, for Acc. feminine “u” - it sounds like “oo” in “going to-oo-oo”. P2019 said that they would not want a step-by-step guide, as they felt that there would be too many things “to keep in your memory”. However, they thought that some kind of very general strategy could be useful. P2001 and P2004 were keen on having “some kind of steps that could be followed to decide on the noun suffix”. It appears that trying to produce case-forms “from memory” was not an effective strategy for my interviewees.

By contrast, P2008 and P2027 were very clear on what they do to get a suffix, though their priorities in the Questionnaire were slightly different, as P2008 starts with a verb, while P2027 first thinks of the context, but both of them have a sequence that they follow. P2008 is convinced that “to learn a case, you need an exemplar phrase in your mind”. This is echoed in the discussion with P2027, who considers applying the suffix to new words very important for language learning but says that, initially, they always come back to the original phrase that they learn with that suffix, but then eventually by-pass that step and get quicker. They expand this idea by saying that they see these initial examples as “structures” where other words could be “slotted in”, or which could be “adjusted or manipulated to suit the context that you need”. However, they felt that they used this strategy considerably less when they

were learning French at school, adhering to pre-learned chunks, as that was a sure way to get a good grade. Furthermore, they added that they “break longer or more complicated sentences into smaller structures, as those are easier to handle”.

Overall, the qualitative data from the present study, show that, in some areas, participants demonstrate very similar attitudes, for example, to learning in general, or to expected difficulties in learning Russian (being reading in Cyrillic), as well as in acknowledging case morphology as most challenging aspect after completing the course. Also, participants’ were unanimous in their perception of the instruction that they received as very positive. On the other hand, there is considerable variation in how participants’ perceive the actual language learning, as for some of them, words and vocabulary appear to be of prime importance, while others concentrate on the structure of the language. These views appear to be reflected in their language learning strategies, which are not always conscious or deliberate. Finally, the top scorers among the interviews appeared to have a well-established strategy for learning case inflection.

6.13. Overall conclusion

It can be concluded that the data collected in the present study provided substantial material for the investigations and detailed analysis enabled me to acquire well-informed answers to the RQs posited. Analysing the data factoring in confounding variables, such as, gender, case contexts and familiarity of lexis has proved to be very beneficial for the current investigation, as that allowed to uncover certain differences, which were not apparent in the initial investigations. Also, the three confounding variables of gender, case contexts and familiarity of lexis demonstrated how the success of case inflection production can be affected at beginner level, and their relationships can be quite complex. Finally, the qualitative analysis of individual success rates provided some invaluable insights into acquisition trajectories for each of the inflections which were investigated. These results will be discussed in Chapter 7.

Chapter 7
DISCUSSION OF THE RESULTS

The main aim of the present study has been to investigate more effective ways of teaching grammar in an L2 classroom through optimizing learners' cognitive processing, focusing more specifically on production of Russian case inflection by English-speaking adult beginners, with no previous knowledge of the language. The research literature was reviewed from two angles – first, studies in L2 teaching and learning, which investigated grammar instruction, were discussed, as I looked at the integration of teaching meaning and form, as well as the role of explicit and implicit (Chapter 2). Second, research on learners' processing and speech production, including the part that WM plays in success of learning, was analyzed (Chapter 3). In addition, I briefly touched upon suggestions of possible application of Skill Acquisition Theory from cognitive psychology, to inflection production, that involving the discussion of proceduralization. Furthermore, I examined how the above issues would be reflected in the production of Russian case inflection specifically (see 4.1). This theoretical background allowed me to draw seven inferences suggesting the ways of optimizing L2 grammar instruction (see 3.4) and to propose my innovative teaching framework implementing those suggestions (see 4.2).

The emperical part of the present study (Chapter 5 and Chapter 6) aimed to test the effectiveness of the proposed framework with regard to two Russian cases, namely, Prep. and Acc. (see 5.1.1 for reasoning), indicated by case inflection accuracy in participants' unprepared oral production after 10 hours of case-learning. At the same time, the same data was analyzed further, in order to investigate the effects of different factors, involved in case inflection production, namely, case contexts, gender and familiarity of inflected nouns, on its success rates. In addition, individual variation in case inflection production, which the proposed framework aims to reduce, was explored from the point of view of the distribution of different success rates, the role of WM and learners' perceptions and attitudes.

In this chapter, the emperical results from Chapter 6 will be discussed in the light of the findings outlined in Chapter 2 and Chapter 3.

7.1. The effectiveness of the proposed teaching framework, addressing learners' processing

At the forefront of my findings is the overall success rate, produced by my beginner participants in the post-test, after just 10 hours of case-learning, which was calculated at

80% (see 6.2.1). This figure on its own demonstrates that the proposed teaching framework, presented in 4.2, can be very effective for teaching Russian cases to English speaking beginner learners, as in SLA, the rate of 60% and over is accepted as the justification of feature acquisition (Vainikka & Young-Scholten, 1994). From the language pedagogy point of view, 80% would be equivalent to a first grade in HE, which also confirms the success of case inflection learning during the teaching intervention.

As explained in 5.2.2, no comparator group was used, due to low RS participant numbers and a number of confounding factors. Furthermore, the results of the current study cannot be compared to those reported in the other studies measuring case inflection success rates, which were identified in 4.1.4, for quite a number of reasons. First, University students in the four studies (see Table 4) and the participants of the present study received very different amount of instruction (between 1645 and 150 hours, as compared to 20 hours in the present intervention). Second, participants in the earlier studies, as well as Prep. and Acc., were producing the other three oblique cases, which would require additional processing resources. In addition, intermediate level participants in Rubinstein's study would also handle more complex vocabulary and syntax. Third, the testing techniques were quite different. The oral test in the form of an interview was only used with more advanced students (Thomson, 1980; Rubinstein, 1995), while Year 1 testees were given a writing task (Arnett & Lysinger, 2013; Cherepovskaia et al., 2021). For the current study, on the contrary, participants supplied oral production samples after 20 hours of instruction. Furthermore, the five-minute traditional Interview part, where participants could use pre-learned chunks or at least forms which have produced before, constituted less than a third (27%) of Test 2 in the current study. The remaining 73% of inflections were produced in completely unprepared sentences during semi-structured elicitation in the specifically-designed Comics Task, which required learners to apply their knowledge of inflection to new contexts, not encountered previously. In addition, in the Comics Test, 18% of inflection contexts involved new items (see 5.4.4 for test design), while no unfamiliar vocabulary was reported in any of the above mentioned studies. Thus, the testing from the previous studies is unable to render results comparable to the present study. For that reason, the results of the present study were assessed against the expected level of acquisition in SLA and the standard HE grading percentages (as explained above).

Comparing the success rates for the two parts of Test 2 demonstrates how testing techniques can affect the results. As expected, the Interview part, in the present study, had a higher success rate of 86% than the overall 80%. This confirms that at least some of the case forms that beginners produce in an interview test are from memory and might not

necessarily demostrate testees' acquisition of grammatical features (see 5.4.4 for discussion). It is further supported by the fact that the success rate in the Interview is even higher than that the overall rate for familiar items (82%), further reinforcing the idea that at least some of the case forms in a language Interview were chunks. Considering that Comics part of Test 2 was quite demanding from the point of view of processing, 78% success rate achieved by the participants, for the cases that they have mastered, despite the absence of motivation to score for grades (see 5.2.1.), appears quite high. This can be considered as even more prominent indication of the effectiveness of the proposed teaching framework from the language pedagogy point of view.

Moreover, delayed testing (Test 3) demonstrated that the case inflection accuracy acquired during the instruction period was retained, as the overall case inflection production success rate was only 2% lower than that in the immediate post-test (Test 2). More so, the range of scores appears to be considerably reduced, which suggests that participants' performance has become more consistent. This further confirms the effectiveness of the proposed framework in establishing strong production skills, including using case inflection in speech.

Furthermore, the qualitative data from Part 1 of the Questionnaire, Learner Diaries and In-depth Interviews demonstrate very positive attitude of learners towards the instruction that they received (see 6.6.2). In their Interview, P2004 gives a clear indication of the spiralling instruction being more effective than the traditional teaching they received at school: "I did 20 lessons with you in Russian and I got to a similar level of speaking that I got to in German, which I did like for three years. I think I found Russian slightly easier to learn". The same participant later on emphasized directly the ease of processing offered by the proposed framework (though this was never articulated to participants): "The way you taught us, helped my brain absorb it better". P2027 focused on the advantage of oral output: "I think the way you taught us is much more efficient, as at school, I knew all the rules, but to produce a verb form, for example, I would need to go through the entire conjugation table to get to the one I needed. In Russian, if we learnt it, I could use it straightaway." P2008: compared the spiralling instruction to repeating after a tape: "so much better to speak about ourselves or something relevant to us, rather than reproducing the phrases mindlessly." This was also confirmed by the anonymous Part 1 of the Questionnaire (see summary of responses in 6.6.2).

Following the above discussion, I believe that the proposed spiralling-based framework, integrating the teaching of grammar with teaching speaking and optimizing learners'

processing, could be considered an effective way of teaching Russian case inflection to complete beginners.

7.2. Processing of case inflection and how it affects the production of Prep. and Acc. at beginner level

7.2.1. Constituent Assembly during inflection production at the onset of learning

Testing Prep. case in isolation in Test 1, before another case was introduced, aimed to examine the case inflection production process at a very early stage of acquisition. The absence of interference from other case inflections, in addition to processing-friendly spiralling, created more favourable processing conditions for inflection production in this round of testing. When analysed from the point of view of the Speech Production model (Bock & Levelt, 1994/2002) (see 3.1), the absence of other cases created a unique situation when the Function Assignment was not an issue, as Prep. was the only case requiring change of inflection in participants' production; thus learners had no problem distinguishing its location function from the Nominative Subject function. Therefore the errors that were produced in Test 1, can be assumed to have occurred during the Constituent Assembly stage of the inflection production (see the adopted Speech Production Model in 3.1). The errors that participants made in Test 1 show the stages of Constituent Assembly, which appears to always start with the base form (see base form error in Example 2).

Example 2.

P2026: *Šokolad – v paket_?*	P2026: Is chocolate in the (plastic) bag [*masc* ***base form***]?
P2012: *Net*.	P2012: No.
P2026: *"E" ..... Šokolad – v paket**e**?*	P2026: "E" … Is chocolate in the (*plastic*) bag [*masc* ***Prep***]?

The example above also illustrates the cognitive nature of inflection processing of inflection at the start of learning, when participant P2026 is still thinking about the inflection after their game partner answers their question. P2026 calls the actual suffix and finally produces the correctly inflected form, while repeating the whole question, when they could have moved on to the next question. It is possible to speculate that adult learners, who study L2 in a formal setting, could be attempting to realize their explicit knowledge of the case inflection in their production. This example also clearly demonstrates that the participant's cognitive processing of inflection is taking longer than communication situation can offer and is lagging behind the inflection production, as suggested by Skehan (1998). This further supports my suggestion that learners' limited WM capacity needs to be addressed by restructuring grammar instruction, especially at beginner level.

Moreover, the above is an excellent example of output modification during peer interaction, discussed in 3.3.3. It occurs in response to internal feedback, as no external feedback is provided (see Swain & Lapkin, 1995). The message is sent for re-processing to the formulation stage (see McDonough, 2005) and is returned with the correct inflected form. This example, which is one out of many, reinforces the importance of peer interaction, which is one of the main principles of the proposed teaching framework, for improving the level of inflection accuracy. Also, see Example 3.

Furthermore, the need for time during inflection production is reflected in the second type of errors recorded in Test 1, which I called "postponed inflection errors" (see 6.1.4 and 6.2.2), when learners manage to produce their inflection, but have a delay in adding it to the base form. This results either in the time gap between the masculine stem and the inflection (e.g., P2017: *Sok – na stakan ... **ne**?* for "Is the juice on the glass … [*masc Prep*]?"), or in some sort of "stacking" of inflections, with feminine base form marker "a" being followed by the Prep. inflection "e"; often in addition to the time gap, but not necessarily (e.g., P2018: *Čaj – v korobka-**e**?* for "Is the tea (bag) in the box[*fem base form*][*fem Prep*]?". As far as my investigations stretch, I was unable to cite this kind of errors in literature. The postponed inflection errors also seem to be pointing us towards the consciousness of adult learners' processing.

Example 3 presents a sequence of sentences with two postponed inflection errors, supplied by P2019 (a few other sentences, which were produced in between by both P2019 and their game partner, are omitted as they did not contain the target noun *sumka* (for "a/the bag")).

Example 3 (P2019)

P2019: *Sok – v sumk**a**?* […]	base form error	P2019: Is the juice in the bag[*fem base form error*]? […]
P2019: *Âbloko – v sumka ...**E**?* […]	postponed inflection error	P2019: Is the apple in the bag[*fem base form*] …[*fem Prep*]? […]
[*long sequence of partner's questions*]	base form error	[*long sequence of partner's questions*]
P2019: *Šokolad – na sumk**a**.* […]		P2019: Is the chocolate on the bag[*fem base form error*]? […]
P2019: *Stakan – v sumk**a** ...**E**.. v sumk**e**?*	postponed inflection attempt with correct outcome	P2019: Is the glass in the bag[*fem base form*] … [*fem Prep inflection*] … in the bag[*fem Prep correct*]?

P2006: *Net*.		P2006: No.
P2019: *Stakan – na sumk**e**?*	Correct	P2019: Is the glass in the bag[*fem Prep correct*]?

If the case forms of *sumka* (for "a/the bag") are followed through the sample, one can see that P2019 starts with a base form error, then produces a postponed inflection error, then after a long sequence of their partner's questions, goes back to producing another base form error. After that, they record postponed inflection as an attempt (before their first fully-formed inflection of "*sumka*") and finally, in the very next question, they deliver a correct case form without any attempts. Similar sequences are observed in other samples (e.g., P2018). This analysis allows me to suggest that, in some cases, postponed inflection could be an intermediate stage in inflection production, which is then usually followed by a correct case form with no attempts. The phenomenon of postponed inflection leads me to infer that adult learners, at the start of their learning, do not tend to encode fully-inflected forms in their memory and that they use their cognitive skills to add inflection to the retrieved base form items.

This inference is also supported by the attempts (which were followed by the correctly inflected forms) recorded in Test 1, which offer a unique insight into learners' thinking process during inflection production. All (but one) attempts include a base form, which, again, indicates that beginners tend to start their inflection process from retrieving the base form of the noun and only then add the inflection to it. E.g., P2028: *Tarelka – na stol … stol**e***? for "Is the plate on the table[*masc base form*] … table[*masc Prep*]?"; P2006: *Čaj – v čaška ...čašk**e***?" for "Is the tea (bag) in the cup[*fem base form*] …. cup[*fem Prep*]?"). It is possible to infer that participants' attempts could represent the first stage of the inflection process, which, in Test 1, happened to be verbalized. The above examples also show that lexical retrieval and case inflection retrieval appear to be two separate processes that take place just before the Constituent Assembly stage, where the actual inflecting process occurs and the case forms are actually put together. In addition, considerable reduction of the proportion of case forms preceded by attempts in Test 2, might indicate that Constituent Assembly was gradually becoming faster. It is possible to suggest that, if learners establish strong retrieval paths during initial Constituent Assembly, this process might become more efficient as learning develops, and possibly more proceduralized.

From the above discussion, it can be stated that Test 1 presented a unique opportunity to examine the Constituent Assembly stage of inflection production, when Function Assignment was fairly straightforward and not taxing on WM, and provided two main outcomes. First, it

demonstrates that, at least on the onset of learning, the oral inflection production for adult learners might be a completely cognitive process. Therefore the teaching of case inflection needs to consider learners' processing limitation for the learning of inflection to be more effective, which is one of the main principles of the proposed spiralling framework, allowing learners to strengthen retrieval paths and Constituent Assembly process. Second, Test 1 provides examples that support the effectiveness of peer interaction in increasing the grammar accuracy in speech specifically for beginner levels, in addition to its well-documented record at higher levels (see 3.3.3). In addition, this also confirms the importance of peer interaction for the proposed framework.

7.2.2. Function Assignment in oral production of Prep. and Acc.

From the language pedagogy point of view, the production of Prep. and Acc. in the present study, was equally successful (with identical success rates of 80%), confirming the effectiveness of the proposed teaching framework for teaching both of these cases. This is in line with the scores for these two cases reported in the four published studies discussed in 4.1.4 (see Table 4). However, from the point of view of processing, the situation appears considerably more complex.

As discussed in 4.1.4, Cherepovskaia et al. (2021) reports that, at A1 level, their participants produced significantly higher rates for Prep. in comparison with Acc., which drop dramatically at A2, and, then, the rates for the two cases level up by B1+ (see Table 5 for details). In the present study, the comparison of Test 1 and Test 2 success rates for Prep. demonstrates a similar pattern. Though the correct score for Prep. in Test 1 was not as high as in Cherepovskaia et al. (85% : 92% respectively), the drop in percentages between Test 1 and Test 2 was not as large either (5% : 39% respectively) (see Table 30). Thus, after the last five hours of learning (from the moment of introduction of Acc. to the end of the instruction period), participants of the current study dropped the success rate for Prep. to 80%, which, serendipitously, is very similar to A2 score calculated for Cherepovskaia et al. (81%).

With regard to Acc. case, Cherepovskaia et al (2021) report a steady increase of its correct production - from 53% at A1 to 90% at C1 (see Table 5). It reaches approximately the same level as Prep. (85%) at B1+ level. In the current study, the percentages of correct Prep. and Acc. inflections level up at the end of the 20-hour instruction period, though at a lower rate of 80% for both cases. Delayed testing (Test 3) in the present study also provided almost

identical rates for the two cases – 77% for Prep. and 78% for Acc.(see Table 39), thus not revealing any significant differences in the ways that the production of these developed.

There are quite a number of reasons that could be suggested for this decrease in accuracy for Prep. The first and foremost is probably the additional step which would logically appear in the production of Prep. inflection with the introduction of Acc., namely, Function Assignment (see the adopted Speech Production Model in 3.1.). Before Acc. is taught (Test 1), the distinction between Nominative Subject function and Prep. location function does not normally present any cognitive challenges to learners, and all the effort is directed to Constituent Assembly (see discussion in 7.2.1), which could explain the higher initial scores. With two oblique cases in production, participants have two sets of choices – one during the Function Assignment stage and the other, still, during the Constituent Assembly stage (discussed in more detail in 7.3.2 and 7.3.3). In addition, in Test 2, the number of Prep. case functions increases, as months and transport vocabulary requiring Prep. are brought into use, with months possibly competing with days of the week, requiring Acc., as both are time references (see Figure 12 for the order of introduction). The influence of other factors (which appears to be crucial) is discussed in the next section.

Furthermore, despite similar rates in Test 1 and Test 2, the development of inflection production was still different for the two cases, as, for example, the variability was reduced only in Acc. (The individual changes in inflection production are discussed further in 7.5.) On the other hand, in Prep., the range of scores remained similar in the two post-tests. Thus the changes that were instigated by the introduction of Acc. (see discussion in 7.2.2), have stabilized considerably after the instruction period was finished. This could possibly be attributed to the psychologists' suggestion that our brain does not stop processing the information, once we stop learning (see incubation period briefly discussed in 4.2.1; also, e.g., Anderson, 2015; Choi & Smith, 2005; Lightbown, 2008). This also means that the distribution of scores grew more similar within the two case sets, which suggests more consistent performance by participants on both cases. Both changes appear positive and further suggest the effectiveness of the proposed framework for teaching Russian case inflection at beginner level.

As there is no further data on the experimental participants' production in the present study, as the teaching intervention was designed for 20 hours (see 5.3.1.), it is not possible to determine how the inflection production would develop, if participants carried on learning. The data demonstrates that learners in the present study have already displayed a trajectory similar to beginners tested by Cherepovskaia et al., but in a considerably shorter period of

time. Thus, as one plausible option, the equal performance on both cases might indicate that the processing-friendly spiralling might have enabled learners to resolve Function Assignment processing challenge in the shorter period of time. Therefore, it is possible to suggest that the proposed framework facilitated the acquisition of Prep. and Acc. cases, leading to more consistent accuracy in inflection production at an earlier level.

Also, it is not unreasonable to speculate that, with the proposed teaching framework, learners could have the potential to carry on improving, applying case inflection to new items, executing both Function Assignment and Constituent Assembly more efficiently and more accurately. In another speculative scenario, participants might undergo some different changes, for example, have a small dip in success rates on previously learned cases every time a new case is introduced, but then overtake the previous scores, as well as successfully producing new case forms. A longitudinal intervention, using the same spiralling principles for teaching all cases or at least the other three oblique cases could make a fascinating research project that might answer this question.

7.2.3. Conclusion

The above discussion suggests that processing plays an exceptionally important role in case inflection production. Using the Speech Production Model (Bock & Levelt, 1994/2002, see 3.1.) has proved to be a useful approach to analysing case inflection production, as it allowed to separate participants' processing challenges in Function Assignment from those in Constituent Assembly.

For example, the increase in cognitive demand is clearly illustrated by the additional step in the production of feminine case forms during Constituent Assembly, producing more errors and attempts. This study reports a different type of error, namely, postponed inflection error, when inflection is added to the base form with a time delay and without taking a feminine base form marker off. In addition, the role of peer interaction in error repair is confirmed for the initial stages of language learning.

Furthermore, the above discussion demonstrates that challenges in Function Assignment are likely to account for considerable differences in the production of the two cases under investigation. Of particular interest, there appears to be a drop in inflection accuracy for Prep. that occurs after the introduction of Acc., and is reported in other studies (see 7.2.1). The influence of specific factors, such as case contexts, gender and familiarity of inflected nouns, discussed in the next section, will contribute further to this discussion.

7.3. Factors influencing case inflection production.

In this section, the differences in inflection production between Prep. and Acc. are discussed through the prism of different factors that can influence the success of the process, namely, case contexts, nominal gender and familiarity of the lexis. These have not been investigated in the studies discussed above.

7.3.1. Distribution of inflections between cases and case contexts

Before discussing the role of the three factors, I will briefly examine the distribution of case forms in beginners' production between the two cases (Prep. and Acc.) and between different case contexts, which reflects some specifics of their production, as well as foregrounding some of the aspects of discussions in the following subsections.

The largest difference in distribution is that Prep. constituted 60% of total case inflections in the post-test (Test 2), that is 1.7 times more than Acc. (716 and 415 respectively). Arnett & Lysinger (2013) and Cherepovskaia et al (2021) provide some statistics on case distribution in participants' writing samples. Though the number of case forms for these two cases for A1 in the study by Cherepovskaia et al., was considerably smaller (146 : 76), the proportion was similar (1.9 times). (According to Cherepovskaia et al., the balance changes in favour of Acc. at higher proficiency levels, which is also confirmed by the data from Rubinstein (1995).) Arnett & Lysinger (2013) also report *more* Prep. inflections in their Year 1 experimental group, though, in their study, the difference with Acc. is smaller.

On the contrary, in the Russian National Corpus (RNC), according to Kopotev (2008), the frequency of Acc. is significantly higher than that of Prep. (27.18% : 12.89%). The imbalance of these two cases at the initial levels might have a few different reasons behind it. Firstly, it is possible that, as Prep. is normally introduced first, learners feel more confident with both Function Assignment and Constituent Assembly for this case, due to more chances for repetitive retrieval (see 3.3.1), thus supply more of it. Secondly, in some situations, say, when using place names, learners could possibly employ avoidance strategies, opting for Prep. over Acc. for the same reason of having well-established retrieval paths for the former. For example, for two panels in the Comics Task, involving a new place name "Novgorod", participants chose to say *On v Novgorod**e*** *[Prep]* (for "He is in Novgorod.") 1.5 times more

often than *On edet v Novgorod_[Acc]*. (for “He is going to Novgorod.”) (72 : 49). However, Arnett & Lysinger found that the proportions of the use of these two cases could vary in different groups. In their Cognitive Grammar group, which focused on explicit grammar explanations, the percentage of Prep. was slightly higher than that of Acc., while traditionally taught students from their control group tended to use more Acc. It is possible to speculate that some of the traditionally taught learners, at the start of learning, might not differentiate between syncretic Nominative masculine and Acc. masculine forms and actively use zero inflection on masculine Acc., unaware of the difference, but still supplying correct scores, though the Function assignment could be faulty. The Comics Test, in the present study, was designed in such a way that suggested the use of the two cases, aiming to control the case distribution. In addition, during Test 2 and Test 3, questions were sometimes asked encouraging learners to use the same place name with both cases, in order to identify whether they differentiate between the two cases (see test design in 5.4.4.). It might be plausible to suggest that learners who are aware of the difference, might be more cautious with using Acc. or decide not too, if unsure, until they get more confident with it. (Also, see further discussion on the Acc. case contexts production in 7.2.4.)

Thirdly, it could be possible to suggest that the more frequent use of Prep. might be explained by the semantics of the nouns that are inflected in different case contexts. None of the four published studies discussed in 4.1.4 (see Table 4), provide a split of production between different case contexts. The analysis of the distribution of produced inflection between different contexts for Prep. and Acc. in the present study (see Table 17 and Table 18) has revealed some interesting tendencies. On one hand, closed lexical sets – months and days of the week – containing a very small number of nouns (e.g., 7 for days of the week) despite high frequency of use, yield rather limited production (11% and 13% of the total respectively). Though months, that require Prep., might have increased the occurrence of the case in the present data slightly (as there are more words for months than for days of the week), that does not appear relevant to Cherepovskaia et al’s study, as the cartoon that they used to elicit written narratives, did not suggest the use of either of these contexts. Nevertheless, their proportion of Prep. at A1 level was still higher than that of Acc.

The production of items belonging to limited lexical sets – musical instruments, games and transport – though has a potential for a wider range, in the present study, delivered a modest 3%, 3% and 5% of total inflections respectively. Several possible reasons could have played a part in reducing their production. Firstly, the musical instruments and games did not figure in the Comics part of the test (in order to keep its length under control), thus were only produced in the Interview part by learners who played something and often one item per

sample. Also, words for transport were introduced at the very end of the instruction period and some learners might not have felt comfortable using these. Finally, there is only one verb that is predominantely used with the first two (*igrat'* for “to play”) and only one for transport (*exat'* for “to travel”), thus restricting the opportunities for sentence building. In any case, these contexts are unlikely to influence the difference in production between the two cases.

Consequently, the main difference in the amounts of inflection production between Prep. and Acc., is between location and direction contexts - 41% and 19% from the total respectively (the 19% atributed to direction within the overall number of inflections, constitutes 46% of all Acc. inflections, which is nearly a half). Thus, Function Assignment for these two contexts would affect the success rates for the two cases most. A possible reason for the larger amount of inflection production for these two contexts could be quite likely due to the fact that the place names, which were inflected in location and direction contexts, belong to open lexical sets that include a wide range of vocabulary. In the Interview, participants actively used names of the places where they or their friends and family lived and worked; that is in addition to new place names, incorporated into the Comics part of the Test.

Finally, it is also possible that the overwhelming dominance of location (Prep. context) might stem from the number of verbs in beginners vocabulary, which could be used with each case. The range of verbs for location is considerably wider (e.g., *žit'* for “to live”, *rabotat'* for “to work”, *igrat'* for “to play”, *pokupat'* for “to buy”, *plavat'* for “swim”), as opposed to two verbs allocated to Acc. in direction context (*idti* for “to go within a walking distance / to walk” and *ehat'* for “to go by transport / to travel”). Therefore, the former provide more opportunities for making sentences within A1 limited language resources, thus more chances for using Prep. in a test, as well as multiple retrieval in practice. Also, in a standard beginners’ curriculum (which was adhered to in the contents for the teaching intervention in the current study), the verbs requiring Prep. are introduced earlier, due to semantics suitable for beginners’ topics, as well as because of belonging to 1st conjugation type, which is taught first. So, by the time of testing at A1 or A2 levels, they would have been used more often in class, learners simply know them better, thus could possibly use them more. Also, verbs of motion used with Acc. are notoriously challenging in Russian. Nevertheless, in the Comics Test the share of the direction context requiring Acc. (see Table 20) is larger than in the Interview part, which could indicate the effectiveness of the Comics test design, preventing participants from avoiding Acc.

The above analysis of the inflection distribution between different case contexts for Prep. and Acc. gives a plausible explanation of why beginner learners in the present study produced more Prep., as well as contributing to the discussion of the initial difference in success rates. It could be possible that more frequent opportunities for using (retrieving) Prep. forms in practice at the onset of learning results in well-established Function Assignment and Constituent Assembly for this case and could be one of the explanations for the above phenomenon. On the other hand, the challenge of inflecting the newly-introduced Acc., in addition to limited use of verbs requiring it, might result in avoidance of it. Considering similar tendencies reported by the previous studies, similar reasons could possibly account for the dominance of Prep. at initial language learning levels in general. Interestingly, a study by Haznedar (2006), who tested nominal inflection production by an English speaking learner of L2 Turkish, also reported the overwhelming dominance of Locative (here the same as Prep.) over Acc., whose suppliance was very poor. So, it might be interesting to investigate the relationship between Prep. and Acc. in other morphologically rich languages. From the language pedagogy point of view, learners' need to be encouraged to use newly-introduced Acc. more and the avoidance of Acc. has to be monitored during oral production activities.

Finally, examining case inflection within different case contexts suggests that its distribution largely depends on the lexical set which is inflected for a particular case, with location and direction contexts, associated with open lexical sets, being most abundant. The influence of case contexts on the success of inflection production is discussed further in 7.3.2.

7.3.2. The effect of case contexts on the success of inflection production: the role of the time of introduction and lexical sets

In addition to the differences in the distribution of case inflection between different case contexts (see 7.2.3), the case production success rates vary considerably between them, as well as within each case, for example, from 68% for direction to 100% for games, both requiring Acc.). As mentioned above, there is no data for specific case contexts in the four studies investigating Russian case inflection production in classroom environment (see 7.1). According to the design of the current study, case contexts differed not only in the lexical sets of the inflected nouns, but also in the time of introduction during the instruction period (see 5.3.2), both appearing to affect the success of inflection production.

As anticipated, the highest accuracy was observed for closed and limited lexical sets, namely, games (Acc.), days of the week (Acc.) and musical instruments (Prep.), where all or the absolute majority of the items were familiar, - 100%, 94% and 93% respectively. Function Assignment for these sets was connected to the semantics of the inflected nouns, for example, musical instruments required Prep., while games needed Acc., both used after the same verb *igrat'* (for "to play"), making learners' choice fairly straightforward, when their explicit knowledge could be applied directly. Moreover, games and musical instruments only figured in the Interview, which means that they were used by those participants who played something, thus are quite likely to have become pre-learned chunks through multiple retrieval. In contrast, the case forms for days of the week were produced in both parts of the test, but significantly more in the Comics part (82%), where the captions on the panels suggested the use of these items, unlike in the Interview part, when they could have been avoided. This might be one of the reasons why this closed set that consists of just seven items and was introduced fairly early in the instruction period, did not supply the highest score; the other probable reason is discussed in 7.3.3. This provides further evidence that language interviews do not necessarily test the application of grammatical rules to new contexts and confirm the effectiveness of the Comics tests designed for the present study for the accuracy of inflection production.

However, the two remaining limited lexical sets, namely, months and transport (both Prep.) provided lower accuracy rates (also quite similar – 77% and 79% respectively), despite both requiring Prep., which overall supplied higher scores in the present study. It looks very likely that the cause for this result lies in the late introduction - both of these contexts were presented to learners at the very end of the instruction period (Weeks 18 and 19). Thus, learners did not have sufficient time to revisit these case forms enough times in order to form chunks or to strengthen the retrieval connection. As a result, inflection in months and transport contexts was acquired not as well as, say, days of the week, which were first encountered in Week 15. Also, it is important to take into account the fact that both months and transport figured predominantly in the Comics part of the Test (94% and 99% of the forms, respectively), which mainly prevented avoidance. The above discussion could be considered as supporting one of the main principles of the proposed teaching framework, suggesting that for the case inflection to be acquired, learners need to frequently retrieve the grammatical feature with new vocabulary items in order to accurately produce the inflected forms in speech, in order to learn to apply inflection to new linguistic contexts, exercising both Function Assignment and Constituent Assembly. Spiralling creates the conditions for this to happen in the classroom environment, thus ensuring the effectiveness of the case inflection acquisition.

As discussed in 7.3.1, the two case contexts that provided most of the inflections in the current study, namely, location (Prep.) and direction (Acc.), were open lexical sets. Thus, they allowed to include new vocabulary items into the Comics part of the test, in order to test participants' Function Assignment on nouns that they have not encountered before, which would inevitably increase the processing load. The two main contexts bore rather different results, unlike the overall rates for the two cases which are identical.

On one hand, the case forms in the location context, which were produced in abundance in both parts of the test, provided the same success rate as Prep. overall (80%), despite including 25% of new place names. This can lead to the conclusion that, with the use of the proposed teaching framework, the transfer of inflection use to new vocabulary can be successful and the Prep. case inflection can be considered acquired after five hours of instruction when Prep. is the only case in the production (but not the only inflection – see Table 9), in addition to five hours of using Prep. together with Acc. Considering that the participants in this study were not learning Russian to achieve a grade or acquire credits, and were not receiving any homework or preparing for tests, this time scale, can be estimated as being in line with a standard University Russian beginners' course, where cases are traditionally introduced one a week, with 3-5 contact hours per week.

On the other hand, the accuracy in the direction (Acc.) context was the lowest among the case contexts (68%), which would still be rated as acquired in SLA (Vainikka & Young-Scholten, 1994), and would be graded as Upper Second 2:1 in HE. This context appears to present most challenges for Russian language learners in this study. Differences in accuracy between location and direction contexts in Russian were also reported by Comer & deBenedette, 2011 (see 4.1.3), though their study produced mixed results. During the present intervention, the directionality was introduced in Week 16, straight after days of the week – thus, learners had reasonable time span to master the inflection in this context. However, the difference in success rates between these two Acc. contexts is very significant (68% and 94% respectively). The most apparent reason is the difference in the type of the lexical sets corresponding to the two contexts – a small closed set of seven days of the week and an open set of common nouns and place names for direction, including new items. Unlike months (the other closed set but with lower accuracy rate), days of the week were in output production for five sessions, which created the potential for establishing strong retrieval paths for this vocabulary pool. It is also possible that some of the inflections were fused with the base form into a chunk, though a small number of base form errors and some attempts in this context indicate the strong retrieval route. Either of these options would

explain high accuracy. The same inflection in the direction context was applied to a considerably wider range of vocabulary, which required the application of inflection rules to new contexts, handling both Function Assignment and Constituent Assembly at the same time, and logically, a longer time span, for the retrieval paths to be established.

In addition, there are some other possible aspects which could have contributed to the difficulties in distinguishing between location and directionality. For example, unlike other contexts, which use dedicated lexical pools (e.g., games, musical instruments), location and direction use the same vocabulary (e.g., *Â student v universitet**e**[Prep.]*. for “I am a student at the University.”; *Â idu v universitet_[Acc.] peškom*. for “I go to University on foot.”). This means that learners cannot assign a case function just by retrieving a lexical item and have to consider a wider context.

Finally, unlike in English, where location and direction are indicated by different prepositions (“at” and “to” in the example above), which clearly indicate the context, in Russian, there are no overt cues, as the same prepositions are used for both contexts, namely, *v* (for “in”, “at” and “to”) and *na* (for “on”, “at” and “on/to”). Thus, to perform Function Assignment, in order to decide on case inflection, learners need to look for other cues. In most cases, that would be the verb (e.g., *idu* for “(I) go” in the example above). This, consequently, would require more processing resources, as learners would need to hold more items in their WM at the same time during the Function Assignment stage. It also needs to be considered that at the same time, other factors would be at play, for example, gender, which is discussed in 7.3.3. in addition, in both Qualitative Questionnaire and In-depth Interviews, participants were unanimous in stating that case inflection is considerably harder than verbal inflection (which was also taught as part of the course). The majority of them could explain the reason and just said that “verb endings” were “so much more straightforward”. It could be possible that the personal pronouns, which have to agree with verbs in Present (the tense taught during the instruction period), serve as distinctive cues for the choice of verbal inflection, while, for case inflection, often functions have to be deducted from semantics of inflected nouns or from the wider context of a sentence.

The above discussion allows me to infer that the success of the production of case inflection, can vary greatly within a case and appear to depend on a number of factors. The comparison of the inflection success rates between different contexts demonstrates that the case contexts influence the inflection accuracy in as much as they determine the type of lexical set that is inflected, which, in turn, is often associated with a particular time of introduction during the course. However, these two are not the only factors that play a part in

the success of case inflection production and the issue is discussed further in the sections below.

7.3.3. The effect of gender on case inflection production

One more factor that could potentially affect the success of case inflection production, which was investigated in the present study (in addition to case contexts, discussed above), is gender of nouns. None of the previous studies discussed above, including those listed in Table 4, examined gender effects in their case scores. In this section, I will discuss the extent of this influence on the production of Prep. and Acc. case inflection in detail.

With regard to distribution of genders among case forms, participants' production in Test 1 was balanced by the sets of objects used. In Test 2, learners' vocabulary production was more diverse, and participants produced more masculine case forms. Unlike the distribution of the two cases (see 7.2.3), this is in line with RNC, RLC and CEFR distribution of genders and was anticipated (see 5.3.2 and 5.4.4). However, the gap in the amount of production between genders is strikingly larger in the Interview part of Test 2, where learners had more choice in what they say than in the Comics part, where panels, to a degree, suggested words that needed to be used. The presence of games context in the Interview part of the test, which provided only masculine items, is unlikely to explain the difference, as months context, also requiring only masculine, balanced it up in Part 2. It could possibly be that the number of English place names where participants lived and worked (which were mainly classed as masculine), might have contributed to the difference. Overall, the distribution of gender in learners' oral production at different levels, especially in spontaneous speech, might be good to be researched in a separate study. Also, in order to ensure that learners get equal amount of practice for inflecting the two genders, it would be beneficial to monitor the amount of feminine nouns in beginners' vocabulary.

In 4.1.2 and 4.2.2, I suggested that inflecting a Russian feminine noun would involve a two-step process during the Constituent Assembly stage, when the overt feminine base form marker (absent in masculine nouns) is removed before a fused gender+case marker is added. My hypothesis that this additional step in feminine case form production puts additional demand for limited WM resources, thus requiring more effort from an L2 learner, is to a degree supported by the Test 1 data, where only Prep. was tested (before Acc. was taught). First, the statistics of the success rates for different genders indicate some difference in the inflection production between genders – participants produced 1.6 times more errors in feminine case forms. Moreover, there are over five times more attempts on

feminine locations than those on masculine, which is a significantly larger difference than that in the distribution of errors. Also, the small number of postponed inflection attempts (see 7.2.1) were recorded exclusively for feminine, with none registered for masculine case forms. Both base form and postponed inflection attempts can be interpreted as an illustration of the additional step in the process of inflecting a feminine noun in Russian.

For this reason, the lower success rates for feminine in Test 2 (where both Prep. and Acc. were produced) were anticipated. Though the actual scores for both masculine and feminine were lower than in Test 1 (in line with the discussion in 7.2.2), the difference in percentages of total correct inflections between genders was quite similar to that in Test 1 (8% as compared to 9% respectively), and though it was not significant, it was nevertheless indicative. However, the situation appeared considerably more complex and further investigations revealed considerable processing challenges, connected with gender.

To start with, the difference in success rates between masculine and feminine in Acc. (which was *not* taught before Test 1) is even larger than that between genders for Prep. in Test 1 (86% and 72% respectively) and *is* significant, which appears to support my suggestion about feminine case forms requiring more processing. However, in Test 2, two-step Constituent Assembly was not the only processing challenge for feminine. Unlike for the majority of Prep. case forms, which are made with the same suffix “e” used for both genders (no subclasses were taught, see 4.1.), in Acc., learners need to choose between *two* suffixes, different for masculine and feminine. This adds yet another step in case form production, this time the choice of the correct gender+case marker during Constituent Assembly. This is in addition to the choice between case functions during Function Assignment (see discussion in 7.2.2). However, my initial suggestion that that might result in lower success rate for Acc. is only confirmed for feminine Acc., as the accuracy for masculine Acc. was, in fact, higher than that for Prep. in Test 2 (86% and 80% respectively).

Then, the success rates for Prep. in Test 2, which are *the same* for both genders (80%), reveal another important role of gender in case production. The decrease in the percentages of correct Prep. inflection from those in Test 1, instigated by the introduction of Acc., was different for different genders - masculine scores dropped by 11%, while the accuracy for feminine came down only by 2%. Thus, it appears that it is the masculine inflection that required more processing effort in this case, which is contrary to my expectations.

As the additional Function Assignment choice presented to learners by the introduction of Acc. (as discussed in 7.2.2) applies to both genders, it cannot explain the gap in success

rate reduction between masculine and feminine. It is the difference in the Constituent Assembly, which was different for the two genders. For masculine nouns, learners had to perform a different procedure in comparison with Test 1 – they needed to decide whether to add the Prep. “e” inflection or to leave the zero-inflection base form as it is, for Acc. (which is syncretic with the masculine base form). This could possibly explain the different patterns in the histograms for masculine Prep. in the two tests (see Figure 32), as well as potentially being responsible for bringing the success rate down by considerably increasing the processing load. In the situation with feminine, the tiny drop of 2% could indicate that the cognitive effort that was required for the two-step inflection of feminine nouns from the very start and initially resulted in the lower accuracy rates for feminine in Test 1, contributed to establishing a stronger Constituent Assembly procedure for Prep. feminine (possibly on the way to proceduralization), which then was not affected much by the introduction of Acc. The histograms for the accuracy for the two feminine suffixes demonstrate similar patterns in the two tests (see Figure 32), supporting this line of thought. It is also possible to speculate that if enough processing resources could be directed to a challenging grammar aspect at the start of learning, then that could potentially result in stronger acquisition of the grammar feature, which is well-retained, thus confirming the effectiveness of the first principle of the proposed teaching framework.

Interestingly, the number of postponed inflection errors, assumed to indicate the increase in cognitive processing demands during inflection production (see 7.2.1), has increased for Prep. masculine from 12 to 18, and reduced in Prep. feminine (from 8 to 5). This change occurred in the same direction as the success rates between Test 1 and Test 2, and might indicate that postponed inflection errors could be an indication of processing overload during Constituent Assembly. However, no further investigations were made, due to a rather small percentage of postponed inflection errors in Test 2 (2% from total inflections and 10 % of all errors).

If the idea of the competition between the two case inflections within each gender is developed further, then one more reason for the drastic decrease in success rate for masculine Prep. could be plausible. In the competition between the overt Prep. “e” and the Acc. masculine zero inflection, the Constituent Assembly of the overt Prep. is likely to require more processing effort than that for Acc. zero inflection. This is also supported by the fact that masculine Acc. success rate is higher than that for masculine Prep. in Test 2 (86% and 80% respectively), despite Acc. function being a new concept for learners. So, Function Assignment for Acc. masculine might not be as much of an issue. If it is, then, assumed that the zero inflection for Acc. masculine is easier to produce than the overt Prep. inflection,

then the significant difference between genders in Acc. can be considerably aided by the absence of overt inflection in masculine, though that is still related to case differences between genders. This is in line with Missing Surface Inflection hypothesis by Prevost & White (2000), discussed in 2.3, suggesting that the overt inflection is more challenging to produce, even if the representation of the grammatical category is formed. As an alternative explanation, it is also possible to suggest that the processing of the overt inflection could be challenging for cognitive resources at the Constituent Assembly stage (and possibly also at the articulation stage) of Bock and Levelt's (1994/2002) Speech Production Model (see 3.1). This would explain why the difference in Acc. scores between genders in Test 2 is larger than the gap between the success rates for masculine and feminine in Prep. in Test 1, where both required an overt marker.

Then, the 100% accuracy in the games context, which is the highest score in Acc., as well as in Test 2, might have also been considerably aided by the fact that the absolute majority of sports games in Russian are masculine, and *all* case forms produced by the participants in a games context, were masculine. Therefore, with no overt inflection and familiar vocabulary, as well as being used in the Interview part only, 100% rate is hardly surprising. It would be interesting to see how masculine and feminine inflection production would compare for a case, where nouns of different genders have overt suffixes that are different, for example, Instrumental (see 4.1.1).

Furthermore, following the same assumption that the overt inflection employs more processing resources and, thus, is harder to produce, the lowest accuracy among the four suffixes that was recorded for feminine Acc., could be attributed to the "u" inflection being overt and new. Consequently, its retrieval during the Constituent Assembly stage of inflection production, would require considerably more cognitive effort than for any of the other three suffixes, as the new retrieval path is still weak and the assembly procedure is not well-established. Examples 4 and 5 vividly illustrate the processing challenges of Constituent Assembly involving overt Acc. feminine "u". Function Assignment for both Prep. and Acc. has already been acquired by both participants (P2001 and P2005), as both choose, more than once, not to use Prep. "e" in Acc. context (while using it correctly in Prep. and producing correct masculine case forms in both cases). At the same time, they fail to add the "u" suffix to feminine place names, though P2005 is clearly aware of it, as they produce it in their attempt for Prep. These data further support the suggestion of overt inflection being more taxing for processing resources.

Example 4 (P2001)

*On idët v Moskv**a**.*	He is going to Moscow*[fem. base form error in Acc.]*
*… muzej v Moskva ... v Moskv**e**.*	The museum is in Moscow*[Prep.fem.base form attempt]* …. in Moscow*[Prep.fem.correct]*
*On idet (v teplohode) v Samar**a** ....*	He is going […] to Samara*[fem. base form error in Acc.]*
*… (v teatre) v Samara .... **E**.*	In the theatre in Samara*[Prep.fem.postponed inflection]*
On idet... v Novgorod_ ... peškom.	He is going to Novgorod*[Acc.masc. correct]* on foot.
*On dumaet ... stul v Novgorod**e**.*	He is thinking (that) the chair is in Novgorod*[Prep.masc. correct]*

Example 5 (P2005)

*A gde stul? - V Novgorod**e**.*	And where is the chair? – in Novgorod*[Prep.masc.correct]*
*… on edet v Novgorod_*	He is going to Novgorod*[Acc.masc.correct]*
*Drug Ivan ponimaet stul (v muzee) v Moskv**u** ... v Moskv**e**.*	The fiend Ivan understands (that) the chair is in Moscow *[Prep.substituted by Acc.attempt]*… in Moscow*[Prep.fem.correct]*
*On edet v Moskv**a**.*	He is going to Moscow *[fem. base form error in Acc.]*
*On edet v aprel'_. – Kuda? – V Samar**a**.*	He is going in April. – Where to? – To Samara*[fem. base form error in Acc.]*
*Gde teatr? - V Samar**e**.*	Where is the theatre? – In Samara*[Prep.fem.correct]*

On the other hand, it is also possible to follow the idea of the competition between two overt inflections for feminine Prep. and feminine Acc. This puts the feminine Acc. score in context of the two competing masculine case forms discussed earlier, justifying the lower rate by both competing feminine inflections being overt, as well as having a three-step assembly process. An important inference can be made from this discussion – that, during the inflection process, the competition could take place between inflections marking different cases for the same gender, for example, masculine Prep. and masculine Acc. This is contrary to Brooks et al. (1993) and Taraban (2004) whose participants supplied mainly correct case marking but wrong gender; that implying that it is the genders that are competing, rather than the cases. However, Kempe & Brooks (2008) later stated that it might be due to the "training regimen" (Kempe & Brooks, 2008: 704). Also, they were testing the inflection acquisition in the artificial language and in laboratory conditions. It is possible that the order and the emphasis of the instruction with regard to gender and case could influence which inflections compete in learners' production. Thus, investigating the relationship between different genders and cases in learners' production when other oblique cases are

introduced, could shed more light on this phenomenon in the future. Also, it would be interesting to see how Russian neuter nouns would behave in the same two cases, as their Acc. form is also syncretic to base form (the same as masculine), but is overt.

Furthermore, the difference in success rates grew even more significant, when correct scores were analysed for each gender within the Acc. main case context – that is, direction (75% for masculine and 57% for feminine). The feminine Acc. score in directionality context is the lowest group score in the data and the only one below expected acquisition level of 60%. At the same time, the accuracy within the days-of-the-week context, also requiring Acc., is 95% and 87% respectively; this is higher not only than the scores in the direction context, but also than those for any context for Prep., which is deemed acquired. This discrepancy between the two Acc. contexts is an interesting one. As days of the week belong to a closed lexical set, there is a high probability that the three feminine and three masculine Acc. forms were acquired by participants as chunks, thus not requiring the effortful Constituent Assembly for this context. This also explains a very high score for Acc. feminine in the Interview (90%), as from 19 Acc. feminine items, 14 were days of the week. On the other hand, directionality was the last Acc. context introduced in the Instruction period, and, consequently, was not in production as long. This shows that both gender and case context impact case inflection production in different ways, and the present study appears to have teased some of these influences apart.

Delayed testing did not produce any significant differences in success rates for the two genders, demonstrating, again, that participants retained their skills of inflecting both masculine and feminine nouns. Similarly to case sets, participants' performance in the two gender sets appeared more similar in Test 3, as the spread of correct scores within the feminine set, where the accuracy was lower, was noticeably reduced. The reduction of variability within the feminine set is a positive development and an improvement in participants' performance.

For each of the four suffixes, on one hand, none of the differences in accuracy between the two post-tests were found significant, thus the case inflection can be considered to be produced with a similar level of accuracy in both tests and the skills retained for all four. On the other hand, the difference between genders in Acc. in Test 3 has grown from 11% in Test 2 to 25% to become significant. This is likely to be the result of two outliers appearing at the lower end of the feminine scores, whose production will be discussed in 7.4. Despite Acc. feminine having the lowest success rate among the four suffixes, with 64% accuracy, it

would be considered acquired by both SLA and language pedagogy (see 7.1), though the directionality context has proved to be the most challenging.

Finally, in one of the In-depth Interviews, when the difference between producing verbal and nominal inflection was discussed, participant P2008 said “because, with verbs, there is no gender” (gender agreement is not required in Present in Russian). It is quite possible that gender is an extra step in Constituent Assembly during case inflection processing, which is missing from the verbal inflection production and makes case inflection more challenging. This adds to the discussion in differences in Function Assignment between the two types of inflection and why participants found verbal inflection so much easier.

The above discussion demonstrates that gender plays a crucial role in the acquisition of Russian cases, as case inflection is acquired differently for different genders. Therefore, it is paramount to ensure that beginner learners have good understanding of the category of gender and can easily differentiate between genders prior to the introduction of the case system. For example, active use of possessive pronouns that agree with nominal gender, in beginners’ production could establish the idea of Constituent Assembly involving gender without Function Assignment. Then, this procedure could be applied to the new concept of case inflection. It appears rather probable that the incomplete acquisition of gender at the start of case learning, would adversely affect case inflection acquisition. This is an important inference for language pedagogy and investigating the effect of enhanced gender instruction on the success of case production could make another interesting study.

In addition, the present study demonstrates that the success rates for feminine case inflection in Russian can be significantly lower than those for masculine, thus might require more focus and more practice in class. The instructional material in Appendix E already contain gradual introduction of case inflection for masculine and feminine. Perhaps, activities involving a choice of case inflection for each gender might contribute to establishing stronger procedures for the two-step Constituent Assembly.

Overall, the role of gender in case acquisition appears to be considerably more complex than initially expected. First, it appears that case inflections compete within each gender, which is contrary to some previous findings. This has a serious implication for teaching Russian case, as the gender has to be acquired well before case is introduced. Second, during inflection production process, gender interacts with a few other factors, for example, case contexts, some of which appear to attract a certain gender. Therefore, some contexts might require more practice than others (e.g., Acc. directionality). Third, feminine case

inflection in both Prep. and Acc. is demonstrated to be more challenging to produce, due to the two-step Constituent Assembly, as well as to the overtness of inflection in Acc., suggesting that practicing inflecting feminine nouns might be beneficial for increasing case accuracy.

7.3.4. Influence of familiarity of lexis on case inflection production

One more aspect of case inflection production which was expected to affect its success, was the familiarity of the nouns that were inflected (see 5.4.1 for discussion). It was expected that success rate for inflecting familiar nouns would be higher than that for producing case forms of rare vocabulary. That was confirmed by the Test 2 data, where the scores for familiar items were marginally higher than those for rare vocabulary. As the knowledge of the vocabulary was not tested, if a participant struggled to produce a word, prompts were available (e.g., via Skype Chat) (see 5.4.4). However, as the data shows, when a word was not frequently used in class (e.g., *teatr* for “theatre”, *teplohod* for a “passenger boat”), testees produced more errors when inflecting it. For example, both items were left in base form by P2004 and were two of their only three Prep. errors in Test 2; and *teplohod* was one of only two items which were not inflected correctly by P2008, from the total of 28 instances of case inflection in their sample.

Though the difference between success rates for familiar and rare items was not significant, it could be inferred that at least some of WM resources were freed, when participants were producing case forms of familiar lexis. However, it did not appear possible to tease apart how much of the difference in scores was due to the ease of lexical retrieval, to stronger inflecting paths or to chunking, or at least that was not feasible within the scale of the present study. It would be fascinating to see whether that could be achieved in a separate experiment. The only indication of chunking in the Interview part of the test is that the success rate for the Interview is higher than that for the familiar vocabulary overall (86% and 82% respectively). Though the difference is not statistically significant, this could be considered indicative that interviews have a considerable amount of chunking. Therefore, testing the application of grammar would be more effective in the new unfamiliar contexts, for example, in picture tasks (see discussion in 5.4.2).

With regard to new items, lexical retrieval per se was not relevant, as the new place names were given as captions on Comics test panels, thus they did not need to be retrieved from memory. Some of the cognitive effort might have been required to pronounce the names written in Cyrillic, but after 20 hours of spiralling instruction, the absolute majority of

participants were confident readers. Then, as both chunking and the strong inflecting paths were naturally excluded from the process, the case forms supplied were due solely to the application of case inflection rules, when participants had to perform both Function Assignment and Constituent Assembly from scratch. The 66% accuracy for new items, though would be considered acceptable in both SLA and language pedagogy, was significantly lower than that for familiar items. This shows that, during the present teaching intervention, participants learned to perform Function Assignment and Constituent Assembly for Prep. and Acc. on the new lexical items (Bruner's (2009/1960) specific transfer) after 10 hours of instruction on the cases, thus clearly confirming their successful acquisition. The lower scores on the new items also demonstrated that applying inflection to unfamiliar items is definitely a more effortful process and is more challenging for learners than producing a previously-used form. This is the aspect of L2 learning that the Skill Acquisition Theory does not appear to be able to address (see discussion in 3.3.2). Creating classroom activities, when learners would be offered to inflect new items in new pragmatic contexts during oral production, should be able to ensure that well established procedures of Function Assignment and Constituent Assembly for case inflection could be applied to the new environment.

One interesting observation appears worth mentioning here – the correct score for months (77%), which was lower than those for other closed and limited lexical sets, due to the late introduction and consequently shorter time span for production, is only slightly higher than the success rate for the new vocabulary in Prep. (74%). Also, the accuracy for months was higher than the overall rate for rare vocabulary (72%), as testees did not have to retrieve these from memory, as they were given as captions in the Comics part. It is very likely that during the two sessions, since months were introduced, learners did not manage to establish strong retrieval paths or form chunks (except perhaps *v marte* for "in March" which was given as an exemplar phrase). Thus, participants were processing the majority of months as new items, applying the well-established Prep. inflection, which months required, to them.

Furthermore, calculating success rates separately for new items in the two main case contexts where they were used, namely, location and direction, showed that the transfer skills are stronger for Prep. location (74%) than for Acc. direction (68%). Though the difference is not significant, this points us towards the inference that the level of accurate application of case inflection to new nouns can be different for different cases, at least at the start of language learning. It might be that the accuracy of the application of inflection to new nouns for different cases could grow closer with increasing proficiency.

When the results were refined further, it has become apparent that, in Prep. location context, the scores were higher for both familiar and new vocabulary (82% and 74%), though the difference was similar to the overall percentages for those groups of words. Also, gender had no effect on the difference between inflection success rates for vocabulary of different level of familiarity, with absolutely identical percentages for masculine and feminine. This demonstrated that participants had formed well-established retrieval paths for inflecting nouns in location function, with some possible chunking, as well as developing strong inflection application skills for Prep.

However, the rates for the Acc. direction, contrary to the expectations, display a totally opposite picture, with lower scores for familiar vocabulary, which were also significantly lower than those for Prep. location. A logical explanation that could be offered, comes from the fact, mentioned earlier, that, in beginners' classroom, both location and direction are taught on predominantly the same set of vocabulary, for example:
Masculine: *On student v universitet**e**[Prep.location]*. for "He is a student at University."
On idët v universitet_[Acc.direction]. for "He is going to University."

Feminine: *Ona živët v Moskv**e**[Prep.location]*. for "She lives in Moscow."
*Ona edet v Moskv**u**[Acc.direction]*. for "She is travelling to Moscow."

As Prep. was taught first, learners would have established strong retrieval paths for Prep. inflection and possibly proceduralized the process or even formed some chunks with most relevant or most frequently used items denoting places. When Acc. was introduced and learners needed to add the new inflection to the same items, their WM was quite likely to go that well-established route of Prep. inflection first, often retrieving either a case form or a suffix, which were wrong for Acc. This resulted in a number of substitutions of Acc. by Prep., which is a new type of errors that was absent in Test 1 (see 7.2.1), thus decreasing the inflection accuracy for familiar vocabulary. This clearly illustrates the competition of the two case functions during the Function Assignment stage of speech production (see 3.1). Interestingly, there were hardly any opposite substitutions, that is, of Prep. by Acc. On one hand, none of those could have been recorded for masculine, as it was not possible to differentiate between base form errors and these opposite substitutions, as both would result in zero-inflection forms. On the other hand, a very small number of opposite substitutions in feminine (when Acc. "u" was used in Prep. context) in comparison with those by Prep. for Acc., still indicates that participants more often erroneously followed a well-established Prep. function assignment paths in Acc. contexts. Some participants even produced a very small number of substitution errors in days of the week context, which has a very high accuracy

rate (predominantly for masculine). Due to high success rates, the numbers of errors, especially for particular types of errors, did not allow to carry out any statistically viable investigations.

An overall success rate for new items in Acc. was higher than that for familiar vocabulary in the same case, but lower than the accuracy for new words in Prep. (neither of the differences were significant). This might be interpreted as participants' successful application of the new Acc. rules to the same new place names that they applied their Prep. inflection to. However, the significant difference between genders in inflecting new items for Acc., brings us back to the earlier discussion about differences in production of overt and zero inflections, as the scores for masculine (82%) are significantly higher than those for feminine. They are also higher than accuracy on the new place names in Prep. location and are identical to percentages for familiar vocabulary in that Prep. context. On the contrary, the 50% accuracy for feminine new items in Acc. direction is the lowest success rate in Test 2 and indicates that participants found producing the new overt inflection on the unfamiliar items for the function that needs to be deducted from wider context, most challenging. A higher inflection accuracy for feminine familiar items in this context could be due to the fact that the majority of them were *v Moskv**u**[Acc.direction]* (for "*to* Moscow"), which was used as an exemplar phrase in class. Although, it cannot be regarded as a chunk, as participants successfully produced *v Moskv**e**[Prep.location]* (for "*in* Moscow") in the same test.

From the above discussion, it can be clearly seen that the familiarity of lexical items plays an important part in the production of the case inflection. However, similarly to the other factors discussed earlier (case contexts and gender), the influence of this factor cannot be considered in isolation, as it is impacted by other factors, such as gender and case context. For grammar teaching, these findings suggest that, in order to establish strong production skills (when Constituent Assembly and Function assignment does not require a lot of cognitive effort), learners need to be encouraged to apply inflection to new lexical items, as well as in new contexts. At beginner level, comics and picture tasks seem to be well-suited for this purpose. In addition, to increase case accuracy for Acc., the use of familiar place names in Acc. contexts (directionality) needs to be built into the instructional design.

7.3.5. Conclusion

According to the above discussion, exceptionally complex inflection production process can be affected by various factors, resulting in considerable differences in accuracy between suffixes within one case, even if the suffix is the same for both genders (as in Prep.). The

three factors that were discussed are case contexts, nominal gender and familiarity of lexis. As long as my investigations could stretch, these factors have not been analysed in the research literature in relation to case inflection.

I argue that gender is one of the crucial factors influencing case inflection production (see 7.3.3). In the present study, case inflections appear to compete within each gender, that is, Prep. masculine inflection competes against Acc. masculine, and Prep. feminine – against Acc., feminine. This finding is of great importance for grammar teaching, as it suggests that, to ensure high case accuracy, the gender category needs to be acquired before cases are taught. Furthermore, feminine case inflection, at least within the cases studied here, appears to be more difficult to produce than masculine, due to the interplay of a number of aspects. For example, the two-step Constituent Assembly or overtness of inflection, both requiring more processing resources. This is in addition to smaller number of feminine case forms produced. Therefore, the instructional design might need to consider paying more attention to practicing feminine case forms in the classroom, as well as slightly increasing the number of feminine items used in case contexts.

The influences of the three factors appear to be interconnected, for example, success of inflection production in a particular case context can be due to the fact that all nouns that are used within a particular context are masculine or happen to be familiar to the participant. That is why, instructors need to be aware that the accuracy in different case contexts could be different and, even when particular case inflection is generally considered acquired, using it in a new context might require additional practice activities. Often, some other factors contribute to the inflection success rate that is produced, such as the amount of repetitive retrieval that the inflection has been subject to, associated with the time of introduction within the instruction period and the amount practice provided. This points directly to oral production being essential for the acquisition of grammar and confirms the effectiveness of the proposed framework, alternating explicit grammar explanations and reading with speaking activities.

In addition, the discussion of the role of familiarity of lexis suggests that applying inflection rules to new vocabulary is usually more challenging for learners' processing resulting in less accurate production. However, the interference of other factors is observed here too, as familiar place names extensively used in Prep. location context, deliver quite low accuracy in Acc. direction context. It is quite possibly that learners' brain automatically opts for well-established Prep. retrieval path. Thus, more time would be required for beginners to learn to

accurately inflect familiar items for the new case, for example, in direction context in Acc., when familiar place names are used in a new case contexts and with new suffixes.

The effects of all factors that were discussed above have one more implication – though production success rates for case inflection are a good indication of how well a particular case is acquired, they do not necessarily reflect a full picture, as case accuracy can vary significantly between different genders, case contexts and lexical sets within one case. That is why, any experiment investigating case production, needs to account for the influence of the factors discussed in this section.

7.4. Individual variability and its relationship between WM

7.4.1. Group correlations between participants' success rates and their WM span

The data in the current study shows that the relationship between WM and accuracy of case inflection production is far from being straightforward.

As discussed in 7.2.1, the number of errors and attempts for feminine case inflection in Test 1 indicated that more cognitive processing is required for Constituent Assembly for this gender. Therefore, it was expected that WM span would strongly correlate with feminine scores. Contrary to the expectations, no relationship was found between WM and gender in the first test. However, there was some correlation with the overall success rate, which might possibly indicate that the effectiveness of Constituent Assembly, whether masculine or feminine, to some degree, depended on the capacity of WM.

In the post-test (Test 2), again, contrary to what was anticipated, the relationship between WM and masculine inflection production was established. This could possibly be connected to the drastic drop in accuracy for masculine Prep. case inflection after the introduction of Acc., which could have possibly required more cognitive effort than any other case forms (see discussion in 7.2.2).

Furthermore, the correlation with familiarity of lexis, was not found. Counter to what was expected, correlation with WM was registered only for familiar vocabulary. This could possibly suggest that applying inflection rules to new items might be outside the realm of WM. For example, Culicover & Jackendoff (2006) suggest that combinatorial thinking is crucial for developing meaning. Also, Anggraeni et al (2019) demonstrated that strong combinatorial skills can facilitate learning. It seems not unreasonable to speculate that

development of combinatorial skills could influence the production of grammar forms, as Constituent Assembly normally requires combining lexical items with overt inflection. Other processing characteristics were suggested to be important for the success of learning, for example, Brooks et al. (2006) found that nonverbal IQ plays a crucial role in learning inflectional morphology (also see, Kempe & Brooks, 2008). These suggestions provide ample opportunities for future research.

Moreover, though no strong correlations were found in Test 3 either, its statistics produced some though-provoking results with regard to WM. For example, the correlation between WM and the overall scores went down from moderate to very weak and stopped being significant, in comparison with Test 2. On the other hand, moderate correlation appeared between WM and feminine inflection success rates, though it was not significant. One of the possible causes for this unexpected result might be that the Acc. overt inflection “u”, which was fairly new, appeared more demanding for WM capacity than any of the previous inflections. This appears particularly plausible in the context of accuracy for Acc., as well as for feminine, becoming more consistent (see 7.2.2 and 7.3.3).

This brings us back to the earlier conclusion that there are a number of factors that influence the success of case inflection production, and WM is one of them. Due to the interplay of various factors influencing the inflection production success (see 7.3), the relationship between the inflection production and WM appears to be quite complex. It could be possible that, rather than impacting a particular aspect of inflection acquisition, WM resources are directed to whichever aspect of inflection production is the most challenging at that particular stage of learning – mastering new skill of inflecting a noun in Test 1, wrestling with Function Assignment for masculine Prep. in Test 2, or two overt feminine suffixes competing in Test 3 (as masculine Prep. appeared to have settled by the time of the delayed testing). This is also in line with the conclusion of Juff & Harrington: “WM is not a unitary construct and that its role varies” (Juff & Harrington, 2011: 137). This, however, is open for further investigations. The correlations with WM within some individual scores are discussed in 7.4.2.

7.4.2. Individual variability and cognitive resources

Similarly to the relationship between group case inflection accuracy and WM, the relationship between individual scores and WM in the present study, appears fairly complex. However, individual trajectories, analysed at three different levels – overall success rates, total scores for each case, as well as total scores for each gender set, and finally, scores for

each gender+case suffix, allowed to observe some tendencies in individual development of case inflection production in the present study.

To start with, while the WM is a stable characteristic, individual scores change across the three tests, for example, the case production accuracy of P2029 deteriorated from 100% in Test 1 to 72% in Test 3, while that of P2015 improved from 59% for Test 1 to 90% for Test 3. It might be natural to speculate that the scores that improved between Test 2 and Test 3, could be explained by some learners studying Russian after the instruction period finished. That was one of the variables which was not feasible to control. However, I am aware that P2013, P2029, P2023 and P2027 carried on learning Russian, with the first two dropping the overall inflection accuracy, while the last two increased theirs. Therefore, this variable might not have had the effect that it could have been expected to have.

What is also interesting is that participants' accuracy did not appear to always develop in the same direction between different tests. The amount of gain or drop could vary significantly between participants (from 1% to 26% between two tests), as well as within one participants' trajectory. The only consistent improvers, unexpectedly, were two participants with the lowest WM, P2015 and P2017, having had the lowest scores in Test 1. In fact, they had the largest overall gains across the three tests. This is contrary to some previous research (e.g., see discussion in Mackey et al., 2002) suggesting that grammar skills are retained better by learners with higher WM capacity. It could be possible that the proposed framework provided their low-capacity WM with the opportunity to process grammar information at their own pace and keep making progress in their own time.

Furthermore, inflection accuracy for Prep. and Acc. in the two post-tests (after the initial drop of Prep. scores in Test 2) appears to develop in the same direction for individual participants – either accuracy increased for both cases or reduced for both cases. The change in success rates varied significantly from -32 to +21. Surprisingly, the top gains in Acc. were made by participants with the lowest and the highest WMs (two from each end).

The conclusion that could be suggested is that learners' interlanguages are very dynamic, possibly more dynamic at the start of learning, than it has been demonstrated before, when an inflection success rate can change considerably within a span of five hours of learning. That is why revisiting the same material, which is one of the principles of the proposed teaching framework, can address this issue by providing another chance of working with the same case, but in a different context and possibly at a different level (see 4.2.3).

The most interesting result, with regard to the individual variability is the changes in accuracy for each individual suffix, which is clearly demonstrated by the distribution of top (100%) scores within individual trajectories. It appears that, once participant reaches 100% accuracy for a particular case suffix, they start focusing on another suffix, while the rate for the first one goes down but can start increasing later. For example, P2027, in Test 1, produced 100% of masculine case forms correctly, but only 80% of feminine forms. In Test 2, they appear to work on Prep. feminine and achieved the top accuracy for it, while Prep. masculine dropped to 76%. In Acc., neither masculine, nor feminine reached 100% in Test 2, quite possibly because all the cognitive effort went into Prep. feminine. However, Acc. feminine got to 100% in Test 3, while accuracy for Prep. feminine reduced from 100% down to 86%. At the same time, the scores for Prep. masculine, in Test 3, started improving (reaching 92%), while those for Acc. masculine decreased to 70%. This could be interpreted as supporting the suggestion made in 7.4.1, that L2 learners re-distribute their WM resources, directing them onto those grammar aspects which they find most challenging at that moment.

Similar tendencies can be observed in other individual trajectories. For example, P2002, having produced 100% accuracy in Prep. masculine in Test 1, never reached 100% in Prep. feminine, which stayed at around 80% throughout the three tests. However, they managed to retain quite a high score for Prep. masculine in Test 2 (90%), as well as supplying 92% for Acc. masculine, possibly attempting to negotiate these two case forms. Eventually, in Test 3, they produce 100% of correct forms for Acc. masculine, but drop Prep. masculine down to 67% in that test. That effort in working out the production of masculine case forms in Test 3, also took WM resources away from Acc. feminine, which dropped to 57% in Test 3. They seemed to be focusing on masculine inflection, thus, never reached top scores for feminine for either of the cases. As well as being in line with the discussion above with regard to re-distribution of cognitive resources between particular inflections, this analysis supports my argument that case inflections compete within genders.

One more example of a rather different trajectory. Unlike the previous two participants, P2023 never produced 100% accuracy for Prep. masculine, as they appeared to concentrate on Prep. feminine, for which they retained 100% scores in the first two tests. It is not unreasonable to suggest that they might have decided to meticulously follow instructor's advice to pay more attention to the two-step production of feminine case forms (see 4.1). Prep. feminine accuracy in the end drops dramatically in Test 3, while Acc. feminine gets to 100%. P2023 also manages to get 100% for Acc. masculine, probably at the expense of low scores for Prep. masculine.

Moreover, participants' attempts to produce 100% accuracy for two case suffixes in one test, often resulted in the detrimental reduction of accuracy for another case suffix. For example, P2017, after gradual increase of correct production of masculine case forms for both Prep. and Acc., finally supplied 100% rate for both in Test 3, at the cost of Acc. feminine, which was not produced correctly at all (0%). Similarly, P2018 focused on masculine case forms at the expense of feminine. The last two participants were the outliers which appeared in the feminine data. It also should be noted that P2017 had the lowest WM span in Test 3 group and P2018 reported previous difficulties in learning languages, as well as being a dyscalculia sufferer. Therefore, there might be some processing difficulties that affect their redistribution of cognitive resources. However, P2013 displayed a similar trajectory, when they concentrated on perfecting Prep., both masculine and feminine (getting 100% scores for Prep. feminine in all three tests), but their Acc. scores plummeted down (from 67% to 29% for feminine).

As a result of the above discussion, it could be possible to tentatively suggest that successful language learners are able to relocate their cognitive resources to process new information, inflection in this case, with some reduction in scores for previously learnt suffixes but with no great detriment to the production. Though clear dips are often observed after initial high scores, the drops never seem to be as low, as the brain seem to identify the problem and re-distribute the necessary resources again to address it. On the other hand, weaker learners appear to persistently pursue the familiar path and require time for the new inflection production to get to the same level as the one which was previously learnt. They are the participants who also tend to produce more case substitution mistakes (e.g., substituting Acc. with Prep.). Thus ability to adapt their cognitive processing to change (e.g., introduction of a new case) and adjusting their production, thus developing new paths alongside the old ones (e.g., assigning a new function to the same place names or assembling new overt constituents) might be the key to successful learning of case inflection, and possibly grammar learning in general.

Chapter 8
CONCLUSION

The present study is dedicated to identifying more effective ways of teaching L2 grammar, and Russian case inflection specifically, and is of great importance for the development of L2 teaching. Moreover, its originality lies in investigating the issue from the point of view of learners' processing, which appears to be a new direction in L2 learning and acquisition research. In order to do that, it synthesized the most relevant findings in language pedagogy, SLA and psycholinguistics, focusing on the role of form, meaning and function in L2 learning; that of explicit and implicit aspects of learning, as well as proceduralization of language skills, which is studied by cognitive psychology; this being novel for Russian language teaching. At the same time, the current study advances the fields of SLA and psycholinguistics by providing rich data from Russian, a language that has not received much attention in the literature in these fields. Furthermore, the production of case inflection in learners' speech is examined within classroom environment, which is the aspect of grammar learning that is mainly under-researched, especially at beginner level (see 4.1.4).

8.1. Originality of theoretical analysis of inflection production

The original approach of using the WM model by Baddeley (1992) for the analysis of learners' production of case inflection allowed me to identify repetitive retrieval as a way of optimizing both secure storage of grammar information in LTM and creating strong retrieval paths for that information during oral production (see 3.3.1). Furthermore, two types of retrieval, namely, recognition and recall, were examined in the context of two influential SLA theories, namely, the Theory of Input Processing (VanPatten, 2004) and within the classic claims made by the Output Hypothesis (Swain, 1985) and related work on L2 speech production (see 3.3.3). In order to apply these to grammar learning within the constraints of a standard beginner language course, 'noticing' and peer interaction were selected as two of the main means of increasing the amount of repetitive retrieval of inflection (see 3.3.4), and, consequently, making the production of grammar features more effective.

Furthermore, the original analysis of Russian case inflection system was carried out through the prism of the Speech Production Model by Bock & Levelt (1994/2002) (see 4.1.2 and 4.1.3). This enabled me to tease apart, to a degree, the roles of form and function of case inflection during two of language production stages, namely, Function Assignment and Constituent Assembly, where the retrieval of inflection are supposed to occur. Discussing Russian inflection production challenges for learners' processing during these two stages,

allowed me to identify factors that would potentially influence the success of inflection production in speech.

In addition, the stages of speech production were briefly compared to those of Skill Acquisition Theory, used in cognitive psychology to explain proceduralization of skills. I argue that Skill Acquisition Theory does not attend to Function Assignment stage, where case function is selected, and quite possibly to some aspects of Constituent Assembly (see 3.2). Both of these stages would require application of learnt procedures to new contexts that often are different from when they were initially acquired. Function Assignment requires analysis of semantic and pragmatic categories, as well as sometimes considering wider linguistic contexts, while Constituent Assembly involves other linguistic categories (e.g., gender or phonology) affecting learners' choices. I believe this is why Skill Acquisition Theory would be unable to explain all aspects of language acquisition, with this argument opening a new exciting line of debate and investigations.

The analysis of the learners' production from the angle of cognitive processing presents a fresh perspective on production of speech, and morphology in particular. At the same time, it allowed me to draw seven inferences regarding ways of optimizing the production process in the classroom environment (see 3.4), which were used as a basis for the proposed teaching framework.

8.2. Innovative teaching framework and measure of its effectiveness

The inferences made from the above analysis allowed me to propose an innovative teaching framework, which is based on processing-friendly spiralling principles, put forward by Bruner 1960/2009 but are developed further (see 4.2). The proposed framework:

- answers the call for developing "a cluster of pedagogical principles and practices" in language teaching (Johnstone, 2004: 667);
- utilizes research findings from three different disciplines (language pedagogy, SLA and psycholinguistics);
- integrates various (sometimes opposite) approaches, for example, grammar is taught in tandem with teaching speaking, systematically alternating attention to form and to meaning, as well as incorporating explicit and implicit elements;
- addresses the constraints of a standard language curriculum, thus contributing to bridging the gap between applied linguistics research and classroom practices;

- makes explicit instruction more processable for L2 learners by segmenting it, thus addressing limited WM capacity;
- provides ample opportunities for repetitive recognition and recall of grammar forms through 'noticing' and regular peer interaction, thus optimizing processing conditions;
- is suitable for complete beginner level;
- can be used for teaching other grammatical features, for example, verb inflection or adjective agreement, as well as more complex grammar categories at other levels, such as Participles or Passive Voice.
- would be applicable to teaching other morphologically rich languages, thus it is potentially seminal to the learning and teaching of other L2s.

The effectiveness of the proposed teaching framework was tested with regard to Russian inflection production in a longitudinal teaching intervention designed for the present study. The contents of the course was in line with a standard beginner curriculum, but was restructured according to the principles of the proposed framework. Thus, the outcomes of the present study are generalizable to standard adult Russian beginner courses, including those attended by ab initio students.

One of the most important findings of the present study is that the proposed teaching framework can be very effective for teaching Russian case inflection to complete beginner learners. It was also found equally effective for teaching the two Russian cases, namely Prep. and Acc., as it enabled learners to produce the inflection for the two cases at the acceptable acquisition level after just 21 hours of case-learning. Furthermore, these results allow me to infer that addressing learners' processing needs in grammar instruction could be a way forward in optimizing language teaching in general.

One of the limitations of the present study was recruiting a reasonable size comparator group (5.2.1), as the pool of Russian Studies ab initio students at home university was quite small, while recruiting more participants from other Russian programmes would have increased the number of confounding factors. Perhaps, for the purposes of language pedagogy, replicating this study at a larger University could enable the investigation of the effect of instruction on two or more groups within similar time scales (spiralling versus traditional).

The other limitation is the transition to the online format at the start of COVID-19 lockdown, which could not be controlled. Though the study was successfully completed and, in the

Questionnaire, not many participants stated that this transition affected them a lot, it would be interesting to see whether a format of learning sessions would have an effect on spiralling. For example, two groups could be taught – one online and another in person, for a comparative study.

8.3. Implications from the empirical investigations

The empirical investigations in the present study demonstrated that case inflection production is an hugely complex process, where various aspects are intertwined and contribute to its success or affect it in a negative way, depending on whether they aid or hinder inflection processing. One of the principle empirical contributions of the present study is investigating the influence of three different factors on case inflection production, namely case contexts, nominal gender and familiarity of lexis, which were not previously explored in research literature. These factors appear to compete for learners' cognitive resources, thus affecting accuracy of different case forms in quite different ways.

Nominal gender was demonstrated to be a crucial factor, resulting in considerable, and often significant, differences in accuracy of case production between genders within one case. This means that case suffixes for different genders *within one case* are acquired differently, including instances when a masculine case suffix is acquired, while a feminine suffix (for the same case) is not. Furthermore, data shows that case inflections compete within one gender, for example, learners normally choose between Acc. masculine and Prep. masculine (rather than between Acc. masculine and Acc. feminine).

This changes the way the Russian cases have been investigated, as the accuracy of production for case suffixes of different genders within the same case, in the present study, had rather different inflection accuracy. Therefore, despite almost identical success rates, the two cases investigated (Prep. and Acc.), developed following rather different trajectories. The significant drop in Prep. success rates at the point of introduction of Acc., which was reported in other studies (see 7.2.2) and observed in the present study, was initially explained by the increased demand on leaners' processing resources during the Function Assignment stage of production. However, the decrease in accuracy only occurred for masculine, as it was competing with zero-suffix Acc. masculine, which had a considerably higher success rate. It would be interesting to see if a similar drop in accuracy would occur in other cases at the point of the introduction of a new case, with two overt suffixes. In addition, the high accuracy for Acc. masculine compensated low scores for Acc. feminine, providing overall success rate for Acc. the same as for Prep., where, in Test 2, both genders had

similar rates. The present study convincingly demonstrated that factoring gender in during any investigation of case is of paramount importance.

This advances our understanding of inflection production process and has important implications for future research. For example, separate studies could investigate the effect of gender in other Russian cases; or comparing production of neuter case forms to that of masculine and/or feminine; or the role of gender in the production of verbal inflection in Past Tense in Russian. Placing this in a broader context, a cross-linguistic study investigating differences between genders in the production of case forms in different Slavic languages or comparing how gender affects production of cases formed with the help of inflection and those expressed by articles (e.g., German) present fascinating opportunities for research. This also has significant implications for the teaching of case inflection, as gender needs to be viewed as an essential pre-requisite for the introduction of cases. Therefore, the acquisition of the nominal gender category has to be ensured before cases are introduced.

In addition, feminine forms were found more challenging to produce from the processing point of view (see 4.1.2). One of the reasons that is suggested, is the two-step inflection process, consisting of the removal of the feminine base-form marker and adding a fused case+gender marker (the first step is not required for masculine, which has no overt inflection in the base form in Russian). In addition, other factors play part in inflection production, such as the overtness of inflection, which was shown by previous research to be harder to produce (see 2.3). Thus, Acc. feminine (requiring a new overt suffix) has provided the lowest scores in this study. This finding regarding difficulties in producing feminine case forms, is of great importance for teaching Russian at beginner level, as it suggests that, in order to increase the accuracy of case production, feminine case forms would require more practice.

Furthermore, familiarity of lexis was expected to ease the processing of inflection, which was mainly confirmed, except for the direction context (Acc.), where the accuracy for familiar vocabulary was quite low. It appears that, as the same vocabulary was used for practicing Prep. location, the retrieval paths were harder to be formed for the new context, as they might have been already established (possibly proceduralized) for Prep. (see 7.3.4). This provides a plausible explanation for difficulties that students of Russian traditionally experience in these two contexts (see discussion in Comer & deBenedette's; 2011).

With regard to case contexts, their effect on case inflection production appears to be mediated by other factors, such as gender of inflected nouns, lexical sets that they belong to

and time of introduction within the instruction period. The novel data in the present study showed that the difference in inflection accuracy between different case contexts of the same case can vary significantly (for example, days of the week with very high accuracy and direction with the lowest among the contexts, both requiring Acc.). From the future research perspective, when case production is investigated, it is important to control or monitor in which case contexts inflection is supplied, as closed and limited lexical sets, for example, are more likely to provide higher success rates, then the open sets; similarly to the contexts requiring only or predominantly masculine nouns. From the point of view of Russian language teaching, direction context, shown to be the most challenging for leaners, similarly to feminine case forms, requires more attention in the classroom. Also, the amount of repetitive retrieval practice provided for particular case context, was also linked to the success of inflection production, thus emphasising the need for introducing different case contexts (even if the inflection is familiar) into beginner learners' oral production activities in class.

Finally, to collect the quantitative data of inflection production, two new testing instruments that address the problem of testing grammar in speech at very low levels, were designed and are now licensed by the University of Leeds (see 5.4.4). The important advantage of these tests is that they enable to test learners' application of inflection to new contexts and to new vocabulary, which demonstrates the level of acquisition in speech, rather than metalinguistic knowledge, which is normally tested in written grammar tests.

Overall, the present study demonstrated that success of case inflection production is a result of the interplay of the whole range of factors that impact inflection processing, often quite considerably, with gender being one of the most significant. Moreover, one and the same factor could have a positive or a negative effect on production, depending on other circumstances (e.g., familiarity of lexis). Therefore, these factors need to be seriously considered in the teaching of cases, as well as being accounted for during empirical investigations of case inflection.

8.4. Individual variability and WM span

The relationship between WM and success of inflection production also was shown to be fairly complex. Some correlations were found, but with different aspects in different tests (see 6.10 and 7.4.1), which led to the suggestion that WM resources are used more actively for the aspects that learners find most challenging at that moment.

The analysis of participants' individual learning trajectories suggested similar tendencies, despite considerable variability. However, only fine-grained analysis of the individual inflection production at the level of four individual suffixes (see 6.11 and 7.4.2), potentially indicated how learners cognitive resources were redistributed during case learning. The present study puts forward a bold suggestion that successful learners are able to direct their cognitive resources to a new case suffix, withdrawing some of the resources from a previously learnt inflection (which had already reached a high score). This results in a slight drop in the accuracy of the latter, though no drastic decline in the success rate is observed. When the production of the new case suffix reaches the top score, some of the resources often can be allocated back, to repair the drop. This repeats for the next suffix, resulting in the new re-distribution, which varies between different learners. Weaker learners appear to gradually increase the accuracy of the initial case suffix (or sometimes, two initial suffixes) at the expense of the new inflection, producing considerably lower scores for the latter. This suggestion offers a new explanation for how learners operate their cognitive resources during grammar learning, and how this process is different between more successful and weaker learners. This is open to debate and requires further empirical investigations, but nevertheless offers a new insight into an on-going debate in psycholinguistics on learner variability.

Moreover, the success of the above re-distribution does not necessarily depend on the size of WM, though the lowest WM capacity appears to be indicative of lower overall scores. However, learners with the lowest WM span in the present study, were demonstrated to be the most consistent improvers, who did not supply any reasonable drops in scores (though they had the lowest scores at the start). This is contrary to some previous findings and invites further research. Finally, a few other cognitive characteristics have been suggested to influence language learning and might play an important role in inflection production (see discussion in 7.4.1).

8.5. Overall conclusion

In conclusion, the present study has reached its goal of identifying more effective ways of teaching Russian case inflection, by proposing an innovative teaching framework, which was successfully tested and provided very positive results. It demonstrated that, in order to help learners increase the accuracy of their case inflection production, the processability of the instruction needs to be addressed and repetitive grammar form retrieval needs to be built into the instruction. Moreover, the study identified some factors that affect case inflection production, with gender being of crucial importance. Finally, it is suggested that more

successful learners redistribute their cognitive resources between the old and new case forms, increasing their accuracy in the new, while maintaining a reasonable level in the old. All of these findings have large implications for Russian language teaching and language pedagogy in general, as well as for future research in SLA and psycholinguistics. It would be probably fair to say that the present study has made an important contribution to answering the question posited at the start of the thesis - How could teachers teach so that learners learn… more effectively?

BIBLIOGRAPHY

Abraham, W., and Leiss, E. (2012). The case differential: Syntagmatic versus paradigmatic case–its status in synchrony and diachrony. *Transactions of the Philological Society*, *110*(3), 316-341.

Ackerman, F. and Malouf, R. (2016). Implicative relations in word-based morphological systems. In: Hippisley A, Stump G (eds.), *Cambridge Handbook of Morphology*. Cambridge: Cambridge University Press, 272-296.

Ackerman, F., Blevins, J. P. & Malouf, R. (2009). Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In Blevins, J. P. & Blevins, J. (eds.), *Analogy in grammar: Form and acquisition.* Oxford: Oxford University Press, 54–82.

Adair-Hauck, Donato and Cumo-Johanssen (2005). Using a story-based approach to teach grammar. In J.L.Shrum and E.W. Glisan (Eds) *Teachers Handbook: Contextualized Language Instruction*. Boston: Thomson Heinle, (3rd ed.), 189-213

Adams, R., Fujii, A., and Mackey, A. (2005). Research methodology: Qualitative research. *Mind and context in adult second language acquisition: Methods, theory, and practice*, 69-104.

Aizlewood, R. (2007). *Routledge Intensive Russian Course.* Routledge.

Alloway, T. P. (2006). How does working memory work in the classroom?. *Educational Research and reviews*, *1*(4), 134-139.

Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European journal of psychological assessment*, *25*(2), 92-98.

Anderson, J. R. (2015). *Cognitive Psychology and its implications.* Worth Publishers. A Macmillan Education Publishers.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning: An introduction. *Studies in Second Language Acquisition*, *37*(2), 185-196.

Arnett, C., & Lysinger, D. (2013). A Cognitive Grammar Approach to Teaching the Russian Case System. *Russian Language Journal/Русский язык*, *63*, 135-168.

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of educational psychology*, *51*(5), 267.

Babby, L. (1980). The syntax of surface case marking. *Cornell working papers in linguistics*, *1*, 1-32.

Babby, L. (1984). Prepositional quantifiers and the direct case condition in Russian. In M. Flier and R. Brecht (eds.), *Issues in Russian morphosyntax*, Slavica Publishers, Columbus, Ohio.

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, *4*(3), 281-288.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of memory and language*, *56*(1), 65-85.

Baerman, M. (2015) Introduction. *The Oxford handbook of inflection*. Oxford Handbooks in Linguistic.

Bailyn, J. (2004). The case of Q. In O. Arnaudova, W. Browne, M. L. Rivero, & D. Stojanovic (eds.), *Formal Approaches to Slavic Linguistics* (FASL-12). The Ottawa Meeting 2003, Michigan Slavic Materials, 49. Ann Arbor, 1–35.

Baker, M. C., & Vinokurova, N. (2010). Two modalities of case assignment: Case in Sakha. *Natural Language & Linguistic Theory*, *28*(3), 593-642.

Bates, E. and MacWhinney, B. (1989). Functionalism and the competition model. In MacWhinney, B. & Bates, E. *The crosslinguistic study of sentence processing*. New York: Cambridge University Press, 3-76.

Bavelier, D., Newman, A. J., Mukherjee, M., Hauser, P., Kemeny, S., Braun, A., & Boutla, M. (2008). Encoding, rehearsal, and recall in signers and speakers: Shared network but differential engagement. *Cerebral Cortex*, *18*(10), 2263-2274.

Bellucci, G., & Dal Pozzo, L. (2016). Acquisition of locative case markers in Finno-Ugric languages: A case study on Finnish L2. *Quaderni di Linguistica e Studi Orientali*, *2*, 141-174.

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied linguistics*, *3*(3), 181-206.

Bischoff, S., & Jany, C. (2013). *Functional approaches to language*, 248, 1-7, Walter de Gruyter.

Bley-Vroman, R. W., Felix, S. W., & Loup, G. L. (1988). The accessibility of Universal Grammar in adult language learning. Second Language Research, *4*(1), 1-32.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In (ed.) M.A. Gernsbacher, Handbook of Psycholinguistics. San Diego: Academic Press; 945-984.

Bock, K., & Levelt, W. (2002). Language production: Grammatical encoding. In (ed.) Altmann, G.T.M., *Psycholinguistics: Critical concepts in psychology*, *5*, 405-452, Routledge.

Bower, G.H. (2000). A Brief History of Memory Research. In Endel Tulving (ed.), *The Oxford Handbook of Memory*, 3-32, Oxford University Press.

Bowles, A. (2003). Implicit And Explicit Learning of Artificial Grammar Structures and Natural Language Morphology: Implications For Second Language Learning (Dissertation)

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, *24*(1), 61-100.

Brooks, P. & Kempe, V. (2013) Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, 41 (2), 281–296.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & cognition*, *41*(2), 281-296.

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. Journal of Memory and Language, 32, 76–95.

Brooks, P. J., Kempe, V., & Sionov, A. (2006) The role of learner and input variables in learning inflectional morphology. Applied Psycholinguistics, 27, 185–209.

Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research,* 20, 436–458.

Brown, J. W. (1977). Comics in the foreign language classroom: Pedagogical perspectives. *Foreign Language Annals*, *10*(1), 18-25

Brown, L., & Iwasaki, N. (2013). Cross-linguistic Influence in the L2 Acquisition of Korean Case Particles by Japanese-Speaking and English-Speaking Learners: L1-L2 Proximity and Learner Perceptions. *Electronic Journal of Foreign Language Teaching*, *10*(2).

Bruner, J. (1990) *Acts of meaning.* Cambridge, MA: Harvard University Press

Bruner, J. S. (1965). *Man: A course of study* (p.4). Educational Services, Incorporated.

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R., J. Jarvelle, and W. J.M. Levelt (eds.) *The Child's Concept of Language.* New York: Springer-Verlag.

Bruner, J. S. (1984). *In Search of Mind: Essays in Autobiography (Alfred P. Sloan Foundation Series)*. New York: Harper Collins.

Bruner, J. S. (1997). Celebrating divergence: Piaget and Vygotsky. *Human development*, *40*(2), 63-73.

Bruner, J. S. (2009; Copyright 1960). *The process of education.* Cambridge, MA: Harvard University Press.

Butler, Y. G., & Hakuta, K. (2004). Bilingualism and second language acquisition. *The handbook of bilingualism*, 114-144.

Bybee, J. L., & Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure*, 45, John Benjamins Publishing.

Cadavid, C. (2009). Teaching English in primary school through a spiral thematic curriculum. *Íkala*, *8*(1), 81-97.

Cary, S. (2009). *Going graphic: Comics at work in the multilingual classroom.* Portsmouth, NH: Heinemann

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual review of applied linguistics*, *25*, 112-130.

Cherepovskaia N, Slioussar N, Denissenko Denissenko A. Acquisition of the nominal case system in Russian as a second language. *Second Language Research*. Published online February 2, 2021

Chohine, I, Perova-Nouvelot, Yu. and Uetova, E. (2019). Some Russian Prepositional Constructions through Russian Learner Corpus (Unpublished AATSEEL Conference presentation) https://hal.archives-ouvertes.fr/hal-02099331/file/AATSEEL%202019%20New%20Orleans.pdf)

Choi, H., & Smith, S. M. (2005). Incubation and the resolution of tip-of-the-tongue states. *The Journal of General Psychology*, *132*(4), 365-376.

Chomsky, N. (2001). Derivation by phase. In: M. Kenstowicz, (ed.). Ken Hale: A Life in Language. MIT Press, Cambridge, MA, 1–52.

Chomsky, N. (2004). Beyond explanatory adequacy. In: A. Belletti, (ed.) *Structures and Beyond: The Cartography of Syntactic Structures*, 3, 104–131, Oxford University Press, Oxford.

Cohen, A. D. (1998). Strategies in learning and using a second language. London: Longman.

Cohn, N. (2014). Building a better 'comic theory': Shortcomings of theoretical research on comics and how to overcome them. *Studies in comics*, *5*(1), 57-75.

Collins, S. (2019). *Neuroscience for learning and development: How to apply neuroscience and psychology for improved learning and training*, Kogan Page Publishers.

Colman A. M. (2015) *A Dictionary of Psychology,* (4), Oxford University.

Comer, W. J., & deBenedette, L. (2011). Processing Instruction and Russian: Further Evidence Is IN. Foreign Language Annals, 44(4), 646–673.

Conway, Andrew, R.A.; Kane, Michael, J.; Bunting, Michael, F.; Hambrick, D. Zach; Wilhelm, Oliver; Engle, Randall, W. (2005) Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review.* 12 (5), 769-786.

Corder, P.S. (1973). *Introducing Applied Linguistics*. London: Penguin Books.

Cornips, L., & Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research*, *24*(3), 267-295.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, *24*(4), 1158-1170.

Crowley, K. (2022). Schema Theory and Language Acquisition: Justification for the Implementation of a Spiral Syllabus. *The English Teacher*, *51*(1), 1-14.

Cubberley, P. (2002). *Russian: A linguistic introduction*. Cambridge University Press.

Culicover, P. W., & Jackendoff, R. (2006). The simpler syntax hypothesis. *Trends in cognitive sciences*, *10*(9), 413-418.

Culicover, P. W., & Jackendoff, R. (2019). In *Current Approaches to Syntax*. P.W.

Culicover and R. Jackendoff (eds), 4. Simpler Syntax, 97-122, De Gruyter Mouton.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic research*, *20*(6), 445-464.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

De Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language learning*, *46*(3), 529-555.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language learning*, 61(2), 533-568.

DeKeyser, R. (2017). Knowledge and skill in ISLA. In *The Routledge handbook of instructed second language acquisition* (pp. 15-32). Routledge.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in second language acquisition*, *19*(2), 195-221.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language learning*, *55*(S1), 1-25.

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.

DeKeyser, R., VanPatten, B., & Williams, J. (2007). Skill acquisition theory. *Theories in second language acquisition: An introduction*, *97113*.

DeKeyser, R.M. (2015). Skill acquisition theory. In: B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction,* (2), 94–112, New York: Routledge.

Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning*, *66*(1), 159-190.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Doughty, C.J., & Williams, J. (1998) Focus on form in classroom second language acquisition, Cambridge: Cambridge University Press.

Durst-Andersen, P. (1996). Russian case as mood. *Journal of Slavic Linguistics*, 177-273.

Durst-Andersen, P., & Lorentzen, E. (2017). Pure case and prepositional case in Russian. *Russian Linguistics*, *41*(2), 177-221.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, *27*(2), 305-352.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 164– 194.

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics—Introduction to the special issue. *Applied linguistics*, *27*(4), 558-589.

ELLIS, R. (1985) Understanding Second Language Acquisition, Oxford: Oxford University Press

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2005). *Instructed second language acquisition: A literature review*. Wellington, Australia: Research Division, Ministry of Education.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL quarterly*, *40*(1), 83-107.

Ellis, R. (2017). Task-based Language Teaching. In Loewen, S., & Sato, M. (Eds.). *The Routledge handbook of instructed second language acquisition.* Taylor & Francis.

Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research*. Routledge.

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System*, *30*(4), 419-432.

Erlam, R. (2008). What do you researchers know about language teaching? Bridging the gap between SLA research and language pedagogy. *International Journal of Innovation in Language Learning and Teaching*, *2*(3), 253-267.

FAERCH, C. and KASPER, G. (eds) (1987) Introspection in Second Language Research. Clevedon: Multilingual Matters.

Felser, C., & Clahsen, H. (2009). Grammatical Processing of Spoken Language in Child and Adult Language Learners. *Journal of Psycholinguistic Research*, 38(3), 305–319. https://doi.org/10.1007/s10936-009-9104-8

Fotos, S. (2001). “Cognitive Approaches to Grammar Instruction.” In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language*, Third Edition. Boston: *Heinle & Heinle*.

Franks, S. (1995). *Parameters of Slavic morphosyntax*. New York, Oxford.

Fulcher, G. (2015). Assessing second language speaking. *Language Teaching*, *48*(2), 198–216.

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language*, *58*(4), 1095-1111.

Gass, S. M. (1997). Input and interaction in second language acquisition. Mahwah, NJ: Earlbaum.

Gass, S. M. (2015). Comprehensible input and output in classroom interaction. In Markee, N. (ed) *The handbook of classroom discourse and interaction*, Wiley Blackwell, 182-197.

Gass, S. M., & Mackey, A. (2007). Input, interaction and output in SLA. In J. Williams & VanPatten (eds) Theories in SLA.

Gass, S. M., & Mackey, A. (2020). Input, interaction, and output in second language acquisition. In *Theories in second language acquisition*, 192-222, Routledge.

Gass, S. M., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. *Modern Language Journal*, 299-307.

Gass, S., & Selinker, L. (1992). *Language transfer in language learning*. Amsterdam: Benjamins.

Geeraerts, D. (2010). Recontextualizing grammar: Underlying trends in thirty years of Cognitive Linguistics. *TABAKOWSKA, E. et al. Cognitive Linguistics in action: from theory to application and back. Berlin: De Gruyter*, 71-102.

Geisler, M., Kramsch, C., McGinnis, S., Patrikis, P., Pratt, M. L., Ryding, K., & Saussy, H. (2007). Foreign languages and higher education: New structures for a changed world: MLA ad hoc committee on foreign languages. Profession, 234-245.

Giancaspro, D., Halloran, B., & Iverson, M. (2015). Transfer at the initial stages of L3 Brazilian Portuguese: A look at three groups of English/Spanish bilinguals. *Bilingualism: Language and Cognition*, 18(2), 191-207.

Gillham, B. (2000). *Developing a questionnaire*, London: Continuum.

Gillham, B. (2000). *The research interview*, London: Continuum.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological review*, *91*(1), 1.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences*, *5*(6), 236-243.

Goldberg, D., Looney, D., & Lusin, N. (2015). Enrollments in Languages Other than English in United States Institutions of Higher Education, Fall 2013. In *Modern Language Association*. Modern Language Association. 26 Broadway 3rd Floor, New York, NY 10004-1789.

Gor K, Chrabaszcz A, and Cook S (2017) Processing of native and nonnative inflected words: Beyond affix stripping. Journal of Memory and Language, 93, 315–32.

Gor, K., Chrabaszcz, A., & Cook, S. (2019). A case for agreement: Processing of case inflection by early and late learners. *Linguistic Approaches to Bilingualism*, *9*(1), 6-41.

Graham, S. (2011). Comics in the classroom: Something to be taken seriously. *Language Education in Asia*, *2*(1), 92-102.

Grenfell, M., & Harris, V. (1999). Modern languages and learning strategies: In theory and practice, London: Routledge. G

Gurzynski-Weiss, L. & M. Baralt (2015). Does type of modified output correspond to learner noticing of feedback? A closer look in face-to-face and computer-mediated task-based interaction. *Applied Psycholinguistics* 36(6), 1393–1420.

Haight, C. E., Herron, C., & Cole, S. P. (2007). The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom. *Foreign language annals*, *40*(2), 288-310.

Han, Z. (2004). Fossilization: five central issues. *International Journal of Applied Linguistics*, *14*(2), 212-242.

Han, Z., & Finneran, R. (2014). Re-engaging the interface debate: strong, weak, none, or all?. *International Journal of Applied Linguistics*, *24*(3), 370-389.

Harley, B. & M. Swain. (1984).The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper & A. Howatt (eds.), Interlanguage, 291–311, Edinburgh: Edinburgh University Press.

Haspelmath, M. (2009). Terminology of case. In Malchukov, A.L., & Spencer, A. (eds). *The Oxford handbook of case*, 505-517.

Hawkins, R. (2003, September). Representational deficit theories of (adult) SLA: evidence, counterevidence and implications. In *European Second Language Association (EUROSLA) conference, 13*.

Haznedar, B. (2006). Persistent problems with case morphology in L2 acquisition. *Interfaces in multilingualism*, 179-206.

Hedge, T. (2000). Teaching and learning in the language classroom. Oxford, England*, Oxford University Press*.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901-931.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33-56.

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, *32*(2), 277-307.

Howatt, A. (1974). “Linear” and “spiral” syllabuses. In Allen, J. P. B., & Corder, S. (1975). *The Edinburgh Course in Applied Linguistics. Papers in Applied Linguistics*, 3, 2-23, Oxford University Press.

Hulstijn J.H. (2005) Theoretical and Empirical Issues in the Study Implicit and Explicit Second-Language Learning: Introduction. Stud Sec Lang Acq, 27(02). doi:10.1017/S0272263105050084

Hulstijn, J. (1989). Implicit and incidental second language learning: experiments in the processing of natural and partly artificial input. In H. W. Dechert, & M. Raupach (Eds.), *Interlingual processes*, 49-73, Language in performance; No. 1. Narr.

Hulstijn, J. H. (2003). Incidental and intentional learning+ In C. J. Doughty & M. H. Long (Eds), *Handbook of second language acquisition,* 349–381, Oxford: Blackwell

Hulstijn, J. H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning. *Implicit and explicit learning of languages*, *48*, 25-46.

Ismailova, K. E. & Gadalina, I. I. (2014) Comparative analysis of case category of nouns in Russian and Tajik languages and formation of grammatical skills. *Polylinguality and Transcultural Practices*, 12/1.

Isurin, L. (2013). Hits and Misses in Teaching Russian in the US: The Perspectives of Instructors, Students, and Enrollment. *Russian Language Journal/Русский язык*, *63*, 25-49.

Izumi, S., & Bigelow, M. (2000). Does Output Promote Noticing and Second Language Acquisition? *TESOL Quarterly*, *34*(2), 239–278. https://doi.org/10.2307/3587952

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in second Language acquisition*, *21*(3), 421-452.

Jackson, F. H., & Kaplan, M. A. (2001). *Lessons learned from fifty years of theory and practice in government language teaching.* George Town University Round Table on Languages and Linguistics (1999), 71-87.

Jakobson, R. (1984 [1958]). Morfologiceskije nabljudenija nad slavjanskim skloneniem. *Russian and Slavic Grammar: Studies 1931-1981*, 105-134.

Janda, A. L., & Tyers, M. F. (2021). Less is more: why all paradigms are defective, and why that is a good thing. *Corpus linguistics and linguistic theory*, *17*(1), 109-141.

Janda, L. (2007). Inflectional morphology. In Geeraerts, D., & Cuyckens, H. (Eds.). *The Oxford Handbook of Cognitive Linguistics*. Chapter 24, 632-650.

Janda, L. A. (2000) *A cognitive model of the Russian accusative case.* Trudy meždunarodnoj konferencii Kognitivnoe modelirovanie, 4, 20-43.

Janda, L. A. (2002). Concepts of case and time in Slavic. *Glossos*, *3*, 1-22.

Janda, L. A. (2013). *A geography of case semantics: The Czech dative and the Russian instrumental*, 4, Walter de Gruyter.

Janda, L. A., & Clancy, S. J. (2002). *The case book for Russian*, 1, Slavica.

Janesick, V.J. (1999) A Journal About Journal Writing as a Qualitative Research Technique: History, Issues, and Reflections. Qualitative Inquiry, 5(4), 505-524.

Jensen, I. N., Slabakova, R., Westergaard, M., & Lundquist, B. (2020). The Bottleneck Hypothesis in L2 acquisition: L1 Norwegian learners' knowledge of syntax and morphology in L2 English. *Second Language Research*, *36*(1), 3-29.

Jezek E, Hanks P. (2010) What lexical sets tell us about conceptual categories?, *Lexis Journal in English Lexicology*, (4).

Johnson B. (2014) *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, (5), Sage Publications.

Johnson, D., & Houia, W. (2005). The Maori Language Curriculum for Mainstream New Zealand Schools: Spiral Lesson/Lesson Sequence Design. *He Puna Korero: Journal of Maori and Pacific Development*, *6*(2), 42.

Johnson, D., & Houia, W. (2005). The Maori language curriculum for mainstream New Zealand schools: Spiral lesson/lesson sequence design. *He Puna Korero: Journal of Maori and Pacific Development*, *6*(2), 42-66.

Johnston, H. (2012). The Spiral Curriculum. Research into Practice. *Education Partnerships, Inc*, 1-2.

Johnstone, R. (2004). Language teacher education. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics. Oxford, UK: Blackwell Publishing Ltd,* 649–671.

Jones, F. R. (1994) The lone language learner: A diary study, *System, 22*(4), 441-454.

Juffs, A. and Harrington, M. (2011) Working memory in L2 learning. Language Teaching, 44, 137–166.

Karpicke, J. D., & Roediger III, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*(2), 151-162.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), The science of writing: Theories, methods, individual differences and applications (pp.57–71).Mahwah,NJ:Lawrence Erlbaum.

Kempe, V. & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, 20, 543-587.

Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning*, 58, 703–746.

Kirkhart, M. W. (2001). The nature of declarative and nondeclarative knowledge for implicit and explicit learning. *The Journal of General Psychology*, *128*(4), 447-461.

Kopotev MV (2008) K postroeniju častotnoj grammatiki russkogo jazyka: Padežnaja sistema po korpusnym dannym [To the construction of the frequency grammar of the Russian language: The case system according to corpus data]. In: Mustajoki A,

Kopotev MV, Birjulin LA, and Protasova EJu (eds) Slavica Helsingiensia 34. Helsinki: University of Helsinki, pp. 136–151.

Kormos, J. (2014). *Speech production and second language acquisition*. Routledge.

Kormos, J., & Sáfár, A. (2006). The role of working memory in intensive language learning. *University of Pécs RoundTable 2506: Empirical Studies in English Applied Linguistics*, 99.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. (1987). *Principles and Practice in Second Language Acquisition*. Englewood Cliffs, N. J.: Prentice Hall.

Krashen, S. (1992). The input hypothesis: An update. *Linguistics and language pedagogy: The state of the art*, 409-431.

Krashen, S. (1993). *The power of reading: Insights from the research.* Westport, CT: Greenwood.

Kudyma, A.S., Miller, F.J. & Kagan, O.E. (2016) *Beginner's Russian: A Basic Russian Course.* Hippocrene Books. Inc, New York

Langacker, R.W. (2007) Cognitive Grammar. In D. Geeraerts, and H. Cuyckens, (eds) *Cognitive Linguistics*. Oxford University Press, 421-462.

Langran, J. & Veshneva, N. (2008) Ruslan 1: a communicative Russian course. Ruslan Ltd.

Lardiere, D. (1998). Case and tense in the 'fossilized' steady state. *Second language research*, *14*(1), 1-26.

Lardiere, D. (2005). On morphological competence. In *Proceedings of the 7th generative approaches to second language acquisition conference (GASLA 2004)*, 178-192.

Lardiere, D. (2008) Feature-Assembly in Second Language Acquisition. In J. Liceras, H. Zobl & H. Goodluck (Eds.) The role of formal features in second language acquisition, 106–140, New York: Lawrence Erlbaum Associates.

Larsen-Freeman, D. (2001). “Teaching Grammar.” In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language, Third Edition*. Boston: Heinle & Heinle.

Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Heinle & Heinle Pub.

Larsen-Freeman, D. (2009). Teaching and Testing Grammar. *The handbook of language teaching*, 518-542.

Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching*, *48*(2), 263-280.

Lasnik, H. (2008). On the development of Case theory: Triumphs and challenges. *Current Studies in Linguistics Series*, *45*, 17-41.

Lawton, J. T. (1977). The use of advance organizers in the learning and retention of logical operations and social studies concepts. *American Educational Research Journal*, *14*(1), 25-43.

Legate, J. A. (2008). Morphological and abstract case. *Linguistic Inquiry*, *39*(1), 55-101.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. In *Cognitive science* (pp. 375-419). Routledge.

Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Han, Z. (ed) *Understanding second language process*, *27* - 44.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in second language acquisition*, *12*(4), 429-448.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in second language acquisition*, *12*(4), 429-448.

Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd éd.). Oxford: Oxford University Press.

Liu, J. (2004). Effects of Comic Strips on L2 Learners' Reading Comprehension. *TESOL Quarterly*, 38(2), 225–243.

Loewen, S. (2011). Focus on form. In *Handbook of research in second language teaching and learning*, 594-610, Routledge.

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285-329.

Loewen, S., & Sato, M. (Eds.). (2017). *The Routledge handbook of instructed second language acquisition*. Taylor & Francis.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input1. *Applied linguistics*, *4*(2), 126-141.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective.* 2(1), 39–52, Amsterdam, the Netherlands: John Benjamins.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), Handbook of language acquisition: Second language acquisition, 2, 413– 468, San Diego, CA: Academic Press.

Loschky, L. & Bley-Vroman, R. (1993) Grammar and Task-based Learning. In G. Crooks and S. Gass (Eds.) Task and language Learning: integrating theory and practice. Clevedon, Avon: Multilingual Matters, 123-167.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432

Lyster, R. and Sato, M. (2013) Skill acquisition theory and the role of practice in L2 development, in Contemporary Approaches to Second Language Acquisition , (eds P. García Mayo, M. Gutierrez-Mangado, and M. Martínez Adrián), John Benjamins,

Ma, R., & Oxford, R. L. (2014). A diary study focusing on listening and speaking: The evolving interaction of learning styles and learning strategies in a motivated, advanced ESL learner. *System*, *43*, 101-113.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27 , 405– 430

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language learning*, *60*(3), 501-533.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language learning*, *60*(3), 501-533.

Mackey, A., Philip, J., Egi, T., Fujii, A. & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactive feedback and L2 development. In (ed) P.Robertson. *Individual differences and instructed language learning,* (2), 181-210, John Benjamins Publishing

MacWhinney, B. (1997) Second language acquisition and the competition model. *Tutorials in bilingualism: Psycholinguistic perspectives*, 113-142.

MacWhinney, B. (2005). A unified model of language acquisition. *Handbook of bilingualism: Psycholinguistic approaches*, *49-67*.

Maksimenko, O.I. and Belyakov, M.V. (2018) *Carnavalization as a reality of language pedagogy: exemplified by the use of comics in teaching Russian as a foreign language*. (Максименко, О. И. и Беляков, М. В.. Карнавализация как реальность лингводидактики: на примере комиксов по русскому языку как иностранному). The Magic of Innovation: Integrative Trends in Linguistics and Foreign Language Teaching (Conference Proceedings). MGIMO, Moscow, Russia. *In press.*

Malchukov, A. L., & Spencer, A. (2009). Introduction. In *The Oxford handbook of case*. Oxford University Press, 1-10.

Mallia, G. (2007). Learning from the sequence: The use of comics in instruction. ImageTexT: Interdisciplinary Comics Studies, 3(3), 1-10.

Martin, J.R. (2009) Genre and language learning: A social semiotic perspective. *Linguistics and Education*, 20, 10–21

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in second language acquisition*, *34*(3), 379-413.

Martin, M. (1995). *The map and the rope: Finnish nominal inflection as a learning target* (No. 38). University of Jyväskylä.

Martin, M. A. (1978). The application of spiralling to the teaching of grammar. *TESOL quarterly*, 151-161.

McCarthy C (2008) Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, 24, 459-86.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in second language acquisition*, *27*(1), 79-103.

McManus, K. (2019). Awareness of L1 form-meaning mappings can reduce crosslinguistic effects in L2 grammatical learning. *Language awareness*, *28*(2), 114-138.

Meisel, J. M. (2000). Revisiting universal grammar. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, *16*(SPE), 129-140.

Menn, L., Duffield C.J. and Narasimhan, B. (2013). Towards an experimental functional linguistics: Production. In Bischoff, S., & Jany, C. (Eds.). *Functional approaches to language*, *248*, 207-249. Walter de Gruyter

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron*, *20*(3), 445-468.

Mizera, G. J. (2006). *Working memory and L2 oral fluency* (Doctoral dissertation, University of Pittsburgh).

Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12(1), 11-37.

Montrul, S. (2010). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, 26(3), 293-327.

Montrul, S., & Yoon, J. (2009). Putting parameters in their proper place. *Second Language Research*, 25(2), 291–311.

Morton, A., Yakimova, L., & VanPatten, B. (2011). *The role of explicit information in the processing of non-canonical sentences in Russian L2.* Paper presented at the Annual Meeting of the American Association of Teachers of Slavic and East European Languages, Pasadena, CA.

Mučnik, I. P. (1971). *Grammatičeskie kategorii glagola i imeni v sovremennom russkom literaturnom âzyke*. Izdatel'stvo" Nauka", Moscow. (Мучник И.П. Грамматические категории глагола и имени в современном русском литературом языке. М. Наука 1971.)

Murre, J. M., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PloS one*, *10*(7).

Nassaji, H., & Fotos, S. S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.

Shintani, N. (2015) The Effectiveness of Processing Instruction and Production-based Instruction on L2 Grammar Acquisition: A Meta-Analysis, Applied Linguistics *36*(3), 306–325, https://doi.org/10.1093/applin/amu067

Neidle, C. (2012) *The role of case in Russian syntax*, *10*, Springer Science & Business Media.

Nevo, E., & Breznitz, Z. (2014). Effects of working memory and reading acceleration training on improving working memory abilities and reading skills among third graders. *Child Neuropsychology*, *20*(6), 752-765.

Newman, A. J., Bavelier, D., Corina, D., Jezzard, P., & Neville, H. J. (2002). A critical period for right hemisphere recruitment in American Sign Language processing. *Nature neuroscience*, *5*(1), 76.

Newman, A. J., Supalla, T., Hauser, P., Newport, E. L., & Bavelier, D. (2010). Dissociating neural subsystems for grammar by contrasting word order and inflection. *Proceedings of the National Academy of Sciences*, *107*(16), 7539-7544.

Norris, J. and Ortega, L., (2000) Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning*, 50 (2000), pp. 417-528

Norton, B., & Vanderheyden, K. (2004). Comic book culture and second language learners. *Sciences and Humanities Research Council of Canada.*

O'Grady, W. (2008). The emergentist program. *Lingua*, *118*(4), 447-464.

Oberauer, Klaus (2009). "Chapter 2. Design for a Working Memory". The Psychology of Learning and Motivation. 51, 45–100.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning.* Cambridge : Cambridge University Press.

Odlin, T. (Ed.). (1994). *Perspectives on pedagogical grammar*. Cambridge University Press.

Offord, D. (1996). *Using Russian: A Guide to Contemporary Usage.* New York: Cambridge University Press.

Ortega, L. (2014). *Understanding second language acquisition*. Routledge.

Oxford, R. L. (2016). *Teaching and researching language learning strategies: Self-regulation in context*. Taylor & Francis.

Oxford, R.L. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle Publishers.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: Benjamins.

Parker, N.V. (2020) Russian in Plain English: a Very Basic Russian Starter for Complete Beginners. Routledge.

Parodi, T., Schwartz, B. D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals, *De Gruyter Mouton.* vol. 42, no. 3, 669-705

Pavlenko. A. Narrative study: whose story is it, anyway? TESOL Quarterly, 36 (2) (2002), pp. 213-218

Pawlak, M., & Biedroń, A. (2021). Working memory as a factor mediating explicit and implicit knowledge of English grammar. *Annual Review of Applied Linguistics*, 1-8.

Pereltsvaig, A., Lyutikova, E., & Gerasimova, A. (2018). Case marking in Russian eventive nominalizations: inherent vs. dependent case theory. *Russian Linguistics*, *42*(2), 221-236.

Pesetsky, D. M. (2013). *Russian case morphology and the syntactic categories* (Vol. 66). MIT Press.

Philp, J., Adams, R., & Iwashita, N. (2013). *Peer interaction and second language learning*. Routledge.

Philp, J., Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: What factors foster a focus on form?. *Language Awareness*, *19*(4), 261-279.

Pica, T. (1988). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language learning*, *38*(1), 45-73.

Pica, T. (2002). Subject-matter content: How does it assist the interactional and linguistic needs of classroom language learners? *The Modern Language Journal*, 86, 1-19.

Pica, T. (2013). From input, output and comprehension to negotiation, evidence, and attention. *Contemporary approaches to second language acquisition*, *9*, 49.

Pienemann, M., & Kessler, J. U. (2012). Processability theory. In Gass, S.M., and Mackey, A. (eds) *The Routledge handbook of second language acquisition*, Routledge: London, 228-247.

Pili-Moss, D., Brill-Schuetz, K. A., Faretta-Stutenberg, M., & Morgan-Short, K. (2020). Contributions of declarative and procedural memory to accuracy and automatization during second language practice. *Bilingualism: Language and Cognition*, *23*(3), 639-651.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization.

Polinsky, M. (2015). When L1 becomes an L3: Do heritage speakers make better L3 learners? *Bilingualism: Language and Cognition*, 18(2), 163-178.

Prabhu, N. (1987) Second Language Pedagogy. Oxford, Oxford University Press.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second language research*, *16*(2), 103-133.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second language research*, *16*(2), 103-133.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, *60*(4), 437-447.

Richards, J. C. (2002). Accuracy and fluency revisited. In E. Hinkel, & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 35–50). Mahwah, NJ: *Lawrence Erlbaum Associates*.

Richards, Jack C. *Communicative language teaching today*. Singapore: SEAMEO Regional Language Centre, 2005.

Rifkin, B. (2005). A ceiling effect in traditional classroom foreign language instruction: Data from Russian. *The Modern Language Journal*, *89*(1), 3-18.

Robin, R., Evans-Romaine, K., Shatalina, G. & Robin, J. (2007). *Golosa: A Basic Course in Russian, Book One*. Pearson, Prentice Hall.

Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language learning*, *45*(2), 283-331.

Robinson, P. (2008) Cognitive Linguistics, Second Language Acquisition, and L2 Instruction—Issues for Research. In P. Robinson & N.C.Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. London: Routledge.

Rodgers, D. M. (2011). The automatization of verbal morphology in instructed second language acquisition. *IRAL, International Review of Applied Linguistics in Language Teaching*; Heidelberg. Vol. 49, Issue 4, pp. 295-319.

Rothman, J., & Slabakova, R. (2018). The generative approach to SLA and its place in modern second language studies. *Studies in second language acquisition*, *40*(2), 417-442.

Rozengrant, S. F. (1987) Error Patterns in Written Russian. *The Modern Language Journal* 71.2: 138-46.

Rubinstein G (1995a) On acquisition of Russian cases by American classroom learners. *International Review of Applied Linguistics in Language Teaching 33: 10–34.*

Rubinstein G (1995b) On case errors made in oral speech by American learners of Russian. *The Slavic and East European Journal* 39: 408–29.

Russell, V. (2014). A closer look at the output hypothesis: The effect of pushed output on noticing and inductive learning of the Spanish future tense. *Foreign Language Annals*, *47*(1), 25-47.

Ryzhova, N. V. (2008) The problem of choice and arrangement of grammar material in teaching Russian to the non-natives. *Russian Language Studies*, 6 (2), 53-59.

Sabourin, L., Stowe, L. A., & De Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*(1), 1-29.

Savage, K. L., Bitterlin, G., & Price, D. (2010). Grammar matters. *New York: Cambridge UniversityPress*.

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2009). *Psychology*. Macmillan.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 3–32

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research,* 12, 40– 72.

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. *Transfer of learning from a modern multidisciplinary perspective*, *3*, 1-51.

Schwartz, M. and Minkov, M. (2014) "Russian case system acquisition among Russian–Hebrew speaking children." *Journal of Slavic Linguistics*, 51-92.

Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209-232.

Selinker, L., & Lamendella, J. T. (1978). Two perspectives on fossilization in interlanguage learning. *Interlanguage Studies Bulletin*, 143-191.

Selinker, L., Swain, M., & Dumas, G. (1975). The interlanguage hypothesis extended to children 1. *Language learning*, *25*(1), 139-152.

Sharwood-Smith, M. (1981) "Consciousness-raising and the Second Language Learner." Applied Linguistics. 11, 159-69.

Sheen, R. (2005). Focus on forms as a means of improving accurate oral production. In A. Housen 8c M. Pierrard (Eds.)*, Investigations in instructed second language acquisition* (pp. 271-310). *Berlin: Mouton de Gruyter*.

Shintani, N. (2015).The incidental grammar acquisition in focus on form and focus on forms instruction for young beginner learners. *TESOL Quarterly* 49(1), 115–140.

Shintani, N., & Ellis, R. (2010). The incidental acquisition of English plural–s by Japanese children in comprehension-based and production-based lessons: A process-product study. *Studies in second language acquisition*, *32*(4), 607-637.

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language learning*, *63*(2), 296-329.

Shvedova, N. Yu., Arutjunova N. D., Bondarko A. V., Ivanov V.V., Lopatin V. V., Uluxanov I.S., Filin F. I., Maslov Ju. S., and Shmelev, D. N. (1980). *Russkaya Grammatika.* (*Русская грамматика*) (Vol. 1). Moscow: Nauka Pubblishing.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skousen, R. (2002). Issues in analogical modeling. *Skousen et al*, 27-48.

Slabakova, R. (2006). Is there a critical period for semantics? *Second Language Research*, *22*(3), 302-338.

Slabakova, R. (2008). *Meaning in the second language* (Vol. 34). Walter de Gruyter.

Slabakova, R. (2009). Features or parameters: which one makes second language acquisition easier, and more interesting to study? Second Language Research, 25(2), 313–324.

Slabakova, R. (2013). What is easy and what is hard to acquire in a second language. *Contemporary approaches to second language acquisition*, *9*(5), 280-295.

Slabakova, R. (2014). The bottleneck of second language acquisition. *Foreign Language Teaching and Research,* 46(4), 543–559.

Slioussar N and Samojlova M (2015) Častotnosti različnyx grammatičeskix xarakteristik i okončanij u suščestvitel'nyx russkogo jazyka [Frequencies of different grammatical features and inflectional affixes in Russian nouns]. In: Proceedings of the conference ‘Dialogue’. Available at: http://www.dialog-21.ru/digests/dialog2015/materials/pdf/SlioussarNASamoilovaMV.pdf (accessed January 2021).

Smith, S. & Crosbi, E.V. (2002) Rus: a comprehensive course in Russian. Cambridge University Press.

Snellings, P., Van Gelderen, A., & De Glopper, K. (2002). Lexical retrieval: An aspect of fluent second–language production that can be enhanced. *Language Learning*, *52*(4), 723-754.

Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated?. *TESOL quarterly*, *42*(2), 181-207.

Special Eurobarometer 386 report “Europeans and Their Languages”. June 2012 http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf

Spencer, A. (2009). Case as a morphological phenomenon. In Malchukov, A. L., & Spencer, A. (eds). *The Oxford handbook of case*. Oxford University Press, 185-199

Spolsky, B. (1969). Reduced redundancy as a language testing tool.

Squire, L.R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory.* 82 (3): 171–177.

Srivastava, N., & Vul, E. (2017). A simple model of recognition and recall memory. *Proceedings of the 31st International Conference “Advances in Neural Information Processing Systems”*, *30*, 292-300

Stakhnevich, J. (2005) Third language acquisition in immersion: a case study of a bilingual immigrant learner Critical Inquiry in Language Studies: An International Journal, 2 (4), pp. 215-232

Stoll, S. (2015). Inflectional morphology in language acquisition. In The Oxford handbook of inflection.

Suzuki, Y. and DeKeyser, R. (2017). Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research*, 21(2), 166–188

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (eds.), Input in second language acquisition. Rowley, MA: Newbury House, 235–253.

Swain, M. (1995). Three functions of output in second language learning. In (eds.)

Swain, M., Cook, G., Seidlhofer, B. (1995). *Principles and practice in applied linguistics: Studies in honor of H.G. Widdowson*, Oxford University Press US, 125-144.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics*, *16*(3), 371-391.

Taraban, R. (2004). Drawing learners’ attention to syntactic context aids gender-like category induction. Journal of Memory and Language, 51(2), 202–216.

Terrell, T. D. (1991). The role of grammar instruction in a communicative approach. *The modern language journal*, *75*(1), 52-63.

Tinsley, T., & Board, K. (2014). Languages for the future: which languages the UK needs most and why. British Council.

Thomson, I (1980) Acquisition of Cases by Students of Russian: Preliminary investigation. Russian Language Journal, 34, 117, pp. 43-52.

Timberlake, A. (2004). *A reference grammar of Russian*. Cambridge University Press.

Tomasello, M. (2005). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Tschirner, E. (1996). Scope and sequence: Rethinking beginning foreign language instruction. *The Modern Language Journal*, 80(1), 1-14.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. *Mind and context in adult second language acquisition*, 141-178.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*, *114*(1), 104.

Ur, P. (2011) Grammar Teaching: Research, Theory, and Practice. In Hinkel (ed) *Handbook of research in second language teaching and learning*, 2, Routledge, 507-522.

Vainikka, A. and Young-Scholten, M. (1994). Direct access to X' theory: Evidence from Korean and Turkish adults learning German. In T. Hoekstra and B. Schwartz (eds.) *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins. 265-316.

VanPatten, B. (2004). Input processing in second language acquisition. *Processing instruction: Theory, research, and commentary*, *5*, 31.

VanPatten, B. (2015). Input processing in adult SLA. *Theories in second language acquisition: An introduction*, 113-134.

VanPatten, B., Williams, J., Keating, G. D., & Wulff, S. (2020). Introduction: The nature of theories. In *Theories in second language acquisition*, 1-18, New York: Routledge.

VanPatten, B., Williams, J., Rott, S., & Overstreet, M. (2004). Form-meaning connections in second language acquisition. In *Form-meaning connections in second language acquisition*, 10-37, Routledge.

Vassilikopoulou, M., Retalis, S., Nezi, M., & Boloudakis, M. (2011). Pilot use of digital educational comics in language teaching. *Educational Media International*, *48*(2), 115-126.

Vygotsky, L.S. (1996). The problem of teaching and cognitive development of school learners. *Psychology and Education* (Выготский, Л. С. Проблема обучения и умственного развития в школьном возрасте. *Психологическая наука и образование*), *1*(4), 5-18.

Wach, A. (2016) L1-based strategies in learning the grammar of L2 English and L3 Russian by Polish learners. *System*, 10(61), 66-73.

Wade, T. (2010). *A comprehensive Russian grammar* (Vol. 8). John Wiley & Sons.

Ward, D. (1993). A Comprehensive Russian Grammar by Terence Wade. *The Modern Language review*, Vol. 88, No. 2, pp. 534-537

Wessel, J. R., Haider, H., & Rose, M. (2012). The transition from implicit to explicit representations in incidental learning situations: more evidence from high-frequency EEG coupling. *Experimental brain research*, *217*(1), 153-162.

White, L. (1989) *Universal Grammar and Second Language Acquisition* (Vol.1). John Amsterdam/Philadelphia: John Benjamins Publishing.

White, L. (1996). Clitics in L2 French. *Language Acquisition and Language Disorders*, *14*, 335-368.

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: language and cognition*, *6*(2), 129-141.

Whong, M., Gil, K. H., & Marsden, H. (2014). Beyond paradigm: The 'what' and the 'how' of classroom research. *Second Language Research*, *30*(4), 551-568.

Williams, N. (1995). The Comic Book as Course Book: Why and How. The 29th Annual TESOL Convention, Long Beach, CA.

Wong, W. (2004). The nature of processing instruction. *Processing instruction: Theory, research, and commentary*, 33-63.

Wong, W., & Van Patten, B. (2003). The evidence is IN: Drills are OUT. *Foreign language annals*, *36*(3), 403-423.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology*, 17(2), 89-100.

Wright, C. (2013). An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English. *TESOl Quarterly*, *47*(2), 352-374.

Wright, C. (2018). *Research in memory and processing in Second Language Acquisition*. In Wright, Pisk and Young-Scholten (eds). *Mind matters in SLA*. Bristol: Multilingual Matters, 203-219.

Xie, T., Wang, T., Wei, Y., & Ye, E. (2019). Declarative memory affects procedural memory: The role of semantic association and sequence matching. *Psychology of Sport and Exercise*.

Xie, Y., Chen, Y., & Li, M. (2020). Convolution forgetting curve model for repeated learning. In *2020 International Conference on Artificial Intelligence and Education (ICAIE)*, 447-454.

Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing*, *33*(4), 497-528.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied linguistics*, 24(1), 1-27.

Zeng, T. (2017) Teaching the Grammar of Russian Language as Foreign in Russian Textbooks for Beginners. *Russian Linguistic Bulletin*, 4(12), 103-107.

Zhan, L., Guo, D., Chen, G., & Yang, J. (2018). Effects of repetition learning on associative recognition over time: Role of the hippocampus and prefrontal cortex. *Frontiers in human neuroscience*, *12*, 277.

Appendix A

List of Russian beginners' course books evaluated by the author

Textbook	*Author/s*	*Year of publishing*	*Publisher*
Beginner's Russian: A Basic Russian Course	Kudyma, A.S., Miller, F.J. and Kagan, O.E.	2015	Hippocrene Books Inc., New York.
Colloquial Russian (A complete course for beginners), 4th edition	le Fleming, S. and Kay, S. E.	2017	Routledge, London and New York
Golosa (A Basic Course in Russian, Book 1), 5th edition	Robin, R., Evans-Romanie, K. and Shatalina, G.	2011	Pearson Education, New Jersey
Live from Russia. (Russian Stage One), 2nd edition	Lekic, M. D., Davidson D. E. and Gor K. S.	2008	Kendall Hunt Publishing Company; American Council of Teachers of Russian, Dubuque, Iowa
Rus: a comprehensive course in Russian.	Smith, S. & Crosbi, E.V.	2002	Cambridge University Press
Ruslan Russian 1 (A communicative Russian course), 5th Edition	Langran, J. and Veshnyeva, N.	2012	Ruslan Limited, Birmingham

Appendix B

Participant Information Sheet 2

01 October 2019

EFFECTIVE WAYS OF TEACHING RUSSIAN TO ENGLISH SPEAKING LEARNERS

You are being invited to take part in a research project. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to participate. You will be able to withdraw at any time. Thank you for reading this.

Research Project (background, aim and duration)

This research project has been initiated by a teacher of Russian with 20 years of professional experience and is funded by the UK Arts and Humanities Research Council. The new teaching approach was first piloted at the University of Swansea, then successfully tested at the University of Sheffield and now is developed further at the University of Leeds as part of a PhD.

This project aims to investigate more effective ways of teaching Russian at beginner level. What this means for you in a nutshell is having free tuition in Russian from a native speaker and an experienced teacher in exchange for taking part in some anonymous language tests. The whole project will involve a series of various teaching experiments, involving different participants, throughout the next two years, each experiment lasting of at least 20 hours.

Why have I been chosen?

For this part of the project, we are looking for English speakers who are keen to learn Russian but have *NO* knowledge of Russian at all. You are offered this opportunity because you have met this initial criteria and responded to the recruitment email.

It is important that you did NOT learn any Russian before the start of this project, so that everybody in a group has a similar language level. We are hoping to have up to 15 volunteers in a group but there might be several groups running at the same time.

What will happen if I take part?

Those who choose to participate will be assigned to a group randomly. You will need to fill in a small questionnaire, to establish that you match the initial criteria. Then you will have up to 20 weekly sessions over two semesters, learning to read and speak Russian with a professional tutor of Russian. The length of each session will be between 1 and 1.5 hours. The starting time of the sessions will be agreed between the participants and the researcher to suit the availability of both parties. The venue will depend on room availability on the main University campus. There will be NO homework given but there will be suggestions about how you can practice your Russian if you wish to do so. This part of the project will be completed by the end of Semester 2.

After completing the scheduled sessions all of the participants will have a written test and each of the participants will have a speaking test individually. Similar tests might be conducted again in 4 – 8 weeks after the first tests to establish whether acquired skills have a long lasting effect, but there will

be NO teaching sessions in between time. There might also be a small test at the end of Semester 1 to assess the progress made up to that point. The tests are likely to be scheduled on the same days as the weekly sessions for each group but the times will be agreed individually. The tests will be anonymous and their results will not be disclosed to your tutors. Please keep in mind that you will NOT be awarded any credits for these sessions or tests.

After the tests are completed, you will be asked to fill in a small questionnaire asking you about your learning.

Do I have to take part?

Participating in this project is absolutely voluntary, i.e. it is up to you to decide whether to take part or not. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. You *can* still *withdraw at any time* without it affecting any of your regular studies in any way. You do not have to give a reason.

What do I have to do if I participate?

If you agree to take part in this project you will be required to attend all scheduled sessions (unless you decide to withdraw) and treat these as your weekly commitment. Please note that this experiment comprises ALL 20 sessions and does NOT have two separate parts. If you miss a session the researcher will offer you a 15 minutes catch-up session before the following group session. To take advantage of this offer you would need to contact the researcher as soon as possible to agree the time and venue for the catch-up session. Participants are encouraged not to abuse this offer and make every reasonable effort to attend scheduled sessions. If a participant misses two consecutive sessions without catching up or without contacting the researcher to arrange a catch-up, he or she might be withdrawn from the project. If the researcher is unable to conduct a session due to the unforeseen circumstances, you will be informed as soon as possible and the session will be rescheduled to another day or time.

You will have to provide your University email address to the researcher, to be informed of any unexpected changes to the scheduled arrangements, should these occur. The researcher's details are provided at the end of this Information Sheet.

All of the materials will be supplied by the researcher. You will be asked to bring these materials to each session if it is possible. You will be allowed to keep any handouts.

As you are going to be working in a group, you will have to be respectful to the other participants and their learning. Please note that any participant can be withdrawn from the project if his/her behaviour is disruptive to the experiment or disrespectful to other participants.

What are the possible disadvantages and risks of taking part?

There are NO risks identified for this project. Please note that the University does not accept any liability with regard to this project, unless it is negligent.

What are the possible benefits of taking part?

The obvious advantage for you as a student of Russian is that you will have up to 20 hours of free tuition with a native speaker who is also an experienced tutor of Russian as a foreign language. By the end of the experiment, we hope that you would be able to sound anything written in Russian, read and understand simple texts within the topics covered, and to answer and ask simple questions, as well as producing short monologues within the vocabulary and grammar studied. In addition, if you complete the experiment and all the tests required within it, we can provide a letter certifying that.

Will I be recorded, and how will the recorded media be used?

The speaking tests will be recorded on digital media. The recording devices and recording procedures will be in agreement with the requirements of the University of Leeds. No video recordings will be made during this project.

The audio recordings of your tests, made for this project, will be anonymised and used for analysis, illustration in conference presentations and lectures; they could accompany a publication of this research. No other use will be made of them without your written permission, and no one outside the project will be allowed access to the original recordings.

All audio recorded data will be kept on a password-protected computer and/or on the University M-drive with high level of security.

Will my taking part in this project be kept confidential?

All the information that is collected from you during the course of this research will be kept strictly confidential by the data controller, that is, the University of Leeds. Your name and your email address will not be disclosed to any third parties and will be destroyed at the end of the research project. You will not be identified in any reports or publications; any of the data collected for this research is going to be kept anonymous and could not be traced back to you. If you have any concerns about data protection, you would need to contact A.J.Slater at a.j.slater@adm.leeds.ac.uk or on 01133433079.

What type of information will be sought from me and why is the collection of this information relevant for achieving the research project's objectives?

Your name and your email address will be used for keeping track of your attendance and contacting you to inform of schedule changes or similar. Other data collected during this project, for example, your age and your mother tongue, are essential for ensuring that all participants start the experiment under equal conditions. The audio recordings of the tests will be analysed and processed at the University and, hopefully, will contribute to the development of methods of teaching Russian as a foreign language. All of the data would be anonymized and could NOT be traced back to you.

What will happen to the results of the research project?

The final results of this research should be ready by the end of 2021. You will not be identified in the report. Those participants who ask to be provided with the results of the research will be sent an executive summary of the results in everyday language. If the results of the research are published, you would be able to find out where to find these by contacting the researcher in 2021-2022.

Due to the nature of this research it might be that other researchers may find the data collected to be useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way and if you agree, we will ensure that the data collected about you is untraceable back to you before allowing others to use it.

After the project is completed, the data used in this research will be stored at the University of Leeds and will be subject to strict regulations with regard to security and confidentiality.

Who has ethically reviewed the project?

This study has been reviewed and given a favourable opinion by Research Ethics Committee of the University of Leeds on "22" February 2019, ethics reference **FAHC 18-060** .

Contact for further information

If you need more information or have any questions about this research project, please do not hesitate to contact:

Researcher: Natalia V. Parker
Email: mlnvp@leeds.ac.uk
Address: 24 John Street, Penarth, Vale of Glamorgan, CF64 1DN

Tel.: +44 (0)7515487975

Supervisor: Dr. James Wilson
Email: j.a.wilson@leeds.ac.uk
Michael Sadler Building, Room 2.04
Tel.: +44 (0) 0113 343 1912

Appendix C

TEACHING RUSSIAN TO ENGLISH SPEAKERS

PRELIMINARY QUESTIONNAIRE

1. Your name…… 2. Your age ……
2. Which Department are you studying in? ……
3. Your first language (that is the language you spoke from birth)……...

   Are you bilingual? *(give details)* ……..
4. Which country were you born? ……
5. Which language did you speak at school? ……..
6. Which language do/did you speak at home? ……....
7. Do you have a Russian relative whom you communicate with? YES/NO

   If yes, a) how often do you communicate? ……..……

   b) for how long each time? ……..
8. Have you ever been to Russia? YES/NO

   If yes, a) how many times? …….....

   b) when was it the last time? ……..

   c) how long do/did you stay? ……..
9. Did you do Russian at school? YES/NO
10. Have you ever attended a Russian class? YES/NO

    If yes, a) when? ……

    b) for how long? …….
11. Have you ever taught yourself any Russian before? YES/NO

    If yes, a) for how long? ……..

    b) how frequently did you study? ……..

Appendix D

Consent to take part in

EFFECTIVE WAYS OF TEACHING RUSSIAN TO ENGLISH SPEAKING BEGINNERS

	Add your initials next to the statements you agree with
I confirm that I have read and understand the information sheet dated "01" October 2019, explaining the above research project and I have had the opportunity to ask questions about the project.	
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences.	
I give permission for members of the research team to have access to my anonymised data. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research.	
I give my full consent for my audio activities during the speaking tests to be recorded on digital media and processed for the purposes stated in the information sheet.	
I agree for the data collected from me, including the recordings of the tests, to be stored and used in relevant future research in an anonymised form. I agree to my data being shared with other researchers, for example, at the conferences, in publications and via dedicated platforms.	
I understand that relevant sections of the data collected during the study, may be looked at by auditors from the University of Leeds or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	
I agree to take part in the above research project and will inform the lead researcher should my contact details change during the project.	

Name of participant	
Participant's signature	
Date	
Name of lead researcher	Natalia V. Parker
Signature	
Date*	

*To be signed and dated in the presence of the participant.

Appendix E

Handout for Lesson 2-1 (first lesson in the second half of the intervention)

NEW WORD: **В** - *in, to*

As you see Russian «В» stands for both *in* and *to* - Russians have a different way of distinguishing between the two functions. Can you work it out if I give you these two phrases: в Омск *(to Omsk)* - в Омске *(in Omsk)*?

WORTH REMEMBERING: *To say that* ***somebody or something is in some place****, for example, in London, you will need* «В» *in front of* Ло́ндон, *as well as* «Е» *attached to the end of the place. E.g.:* **в** Ло́ндоне - ***in*** *London. This works for most places.*

To make it clear: Remember to read **«Е» at the end of places**. As always, «Е» would indicate that the consonant before it is palatalised. What is most important though, is to keep the stress in place, e.g.: Но́вгород - в Но́вгороде. Remember that if «Е» is away from the stress, it does not need too much effort, you still need to read it though.

Ex.3 a) Reading for meaning. Translate each small phrase you read. The names with «Е» at the end should appear without it in English. E.g.: в Омске – in Omsk_.

в Иркутск - в Иркутске	в Ростове - в Ростов	в Магадане	в Мадриде
в Смоленск - в Смоленске	в Саратове - в Саратов	в Воронеже	в Берлине
в Норильск - в Норильске	в Тамбове - в Тамбов	в Хабаровске	в Вашингтоне

Check *that you remember that Russians read* «В» *at the end as ff in Smirnoff, but before* «Е» *it goes back to normal /ъ/ but palatalised.*

b) Big sort out. Look at the list below. Read out only those phrases where «В» is translated as *to* and then only those where «В» stands for *in*. Make sure you read «Е» at the end (if you have «Е»).

в банк, в Ло́ндоне, в па́рке, в магази́н, в колле́дж, в кла́ссе, в университе́т, в рестора́не, в теа́тре

NEW WORDS: **ГО́РОД** – *city, town*
ДЕРЕ́ВНЯ - *a village, (often) countryside*

Check *that all three consonants in your* дере́вня *are palatalised and that the first «е» is weak;*

c) Reading for meaning:

Это го́род Санкт-Петербу́рг, а это река́ Нева́.
А в го́роде Смоле́нске - река́ Днепр. А это мост.
Одо́ев - это го́род или дере́вня? - Го́род.

А Апу́хтино – это дере́вня? – Да, дере́вня.
Я завтра в го́роде. А ты где? – Я дома.
У меня́ есть карта. Где твоя́ дере́вня? – Вот.

WORTH REMEMBERING: *Some small Russian words with only one vowel in them (“****monosyllabics****”)often behave differently and sometimes even have a different ending, e.g.:* **САД – В САДУ́** *(in the garden). They are normally the words which do not sound like English words - they are of Old Russian origin. The words like* парк *or* банк *(which are similar to English) go with the general rule, e.g.:* в па́рке, в ба́нке. *There are not that many of the words like* сад, *so we will be learning these as we go.*

Ex.4 a) Let’s get this right. Put «В» in front of the words below and add the ending, keeping the one stress exactly where it was. Find one word which would need «У» at the end (instead of the normal «Е»), as well as shifting the stress. Keep track of what it is that you are saying.

Tip! *The word we’ve learnt before for “at home”* (до́ма) *is a thing of its own. It doesn’t need anything before it or any changes at the end. However, you might come across the other phrase* «в доме», *which stands for “in a/the house” (as opposed to the outside), it does not necessarily refer to your own house.*

класс, го́род, магази́н, Манче́стер, сад, рестора́н, Ки́ев, колле́дж, банк, музе́й, Портсмут, университе́т, дом,

b) Reading for meaning. As well as understanding what you are reading, try to remember the information too – after all we often need to do that in real life.

Твоя́ сестра́ дома? - Нет. Она́ в колле́дже.	Мо́жно хлеб? – Да, но извини́: хлеб в паке́те.
А твой друг где? – В го́роде Ку́рске.	Извини́те. Что это в стака́не? Лимона́д? – Нет, сок.
Ба́бушка в до́ме? – Нет, она́ в саду́.	У тебя́ ручка есть? – Да, в столе́. - Спаси́бо.

***Check** that you understand that* «в столе́» *means "inside the desk" (in a drawer); that you noticed that the stress in* «в столе́» *is at the very end.*

c) Let's try it in Russian. Try answering the following questions in Russian without looking at the exercise above - you can check if you forget the information but answer without reading. Speaking and reading are different language skills!

***Tip!** In this exercise you would NOT need words like* мой / твой *as Russians do not use them as much as people do in English.*

Кто в Ку́рске?	Ба́бушка в саду́ или в до́ме?	Сестра́ дома?
Где хлеб?	Что в стака́не?	В столе́ ру́чка или каранда́ш?

Have you realised that in this exercise you got the information in Russian, understood the question about it in Russian and then answered it in Russian? How good is that?!

d) Now it's your turn. Make a list of about five people you know and a separate list of where they are, in a different order (!). If you have time to write all the names in Russian, that would be great, but if you are not learning to write, pressed for time or do not have all the letters that you need yet, you can leave them in English. Make phrases in Russian about each person. If you work with someone else, swap lists and ask questions that require «Да»/«Нет» answers, to find out where your partner's people are. Remember to swap as soon as you get a «Да». E.g.: 1) Дже́нни – в Манче́стере? – Нет. 2) Дже́нни в Бри́столе? – Да, она́ в Бри́столе.

***Tip!** Foreign place names which do not end in a consonant or an «А» do not take any endings in Russian. I guess they cannot be classed as masculine or feminine, thus do not fit into the Russian grammar system. E.g.:* Глазго – в Гла́зго.

***WORTH REMEMBERING**: If a place has a vowel at the end (e.g.,* Москва́, Ива́ново *or* дере́вня*), then this vowel is replaced by* «Е». *So, instead of adding* «Е» *to the end of a word as we did with masculine places (that we have had until just), in words with a vowel (e.g., feminines) we need to **drop this vowel first**. E.g.:* **Москва́ - в Москве́** *or* Ива́ново - в Ива́нове *or* дере́вня - в дере́вне. *Make sure that you **keep the stress** where it was - if the stress needs to move, the books and dictionaries would tell you (e.g.:* в столе́*). Your main job is to keep it.*

Ex.5 a) Reading for meaning. Read each phrase aloud, pronouncing all the «Е»-s (find three words where «Е» is stressed at the end). For each phrase give the dictionary form of the word (with a vowel at the end – «А» / «Я» / «О»). E.g.: в Москве́ *(in Moscow)* - Москва́ *(Moscow)*. You will have one word which would have «О» and another one which would have «Я» at the end.

в ва́нне	в кварти́ре	в ко́мнате	в таре́лке	в дере́вне
в ча́шке	в реке́	в окне́	в су́мке	в Москве́
в шко́ле	в библиоте́ке	в кни́ге	в коро́бке	в Кана́де

***Check** that you understand that* в таре́лке *translates as "in a/the dish" and that* в дере́вне *can be translated as "in a/the village" and as "in the country".*

b) Let's try it in Russian. Let's imagine that this is a list of students with their summer placements. Say in Russian who is where this summer. You would need «В» in front of each place and «Е» at the end. Don't forget to drop the vowels at the end (if you have them). Make sure you do it all aloud.

Я́на Соро́кина - Тобо́льск	Мари́я Рае́вская - Челя́бинск	Зинаи́да Губе́нко - Лиепа́я
Никола́й Шу́хов - Элиста́	Таи́сия Черно́ва - Калу́га	Глеб Ли́син - Па́вловск
Валенти́на Ры́бникова - Я́лта	Генна́дий Пульно́в - Бо́лдино	Па́вел Ермола́ев - Алу́шта
Алекса́ндр Жда́нов - Сара́тов	Оле́г Су́рский - Санкт-Петербу́рг	Ири́на Бобро́ва - Хаба́ровск

Appendix F

Objects for "Guessing game" (Test 1)

masculine		***feminine***		***neuter***
"stakan" "*a glass*"	"šokolad" "*chocolate*"	"korobka" "*a box*"	"sumka" "*a bag*"	"moloko" "*milk*" (a bottle or a carton)
"banan" "*a banana*"	"paket" "*a plastic bag*"	"voda" "*water*" (a water bottle) .	"ručka" "*a pen*"	"âbloko" "*an apple*"
"karandaš" "*a pencil*"	"čaj" "*tea*" (a teabag)	"čaška" "*a cup*"	"tarelka" "*a plate*"	

Appendix G

EFFECTIVE WAYS OF TEACHING RUSSIAN TO ENGLISH SPEAKING BEGINNERS

QUANTITATIVE QUESTIONNAIRE
Part 1
(online, anonymous)

1) **Why did you decide to take on learning Russian?**
(You can choose more than one reason – place a "V" in a box to the right.)

general interest		it's something different		it will be useful *(please comment)*	
always wanted to try		like a challenge		love learning new things	
had a go before and failed		was looking for something to do		other reason *(please comment)*	

Comments, if any:
……..
……
……..

2) **BEFORE you joined our project, did you think that you would find learning Russian difficult?**
YES/NO

(comment)
If you answered "Yes", could you please choose the reason why:
(You can choose more than one reason – place a "V" in a box to the right.)

it is very different from English		I'd be taught in an unusual way		its grammar is difficult	
it has a different alphabet		because of its pronunciation		I don't feel comfortable speaking	
I am not good at languages		it has lots of suffixes		I would make lots of mistakes	

If you answered "No", could you please comment why you did not think that
...……..
……
……

Any other comments
……..

3) DURING your classroom sessions, did you feel that the explanations were:
(Please highlight ANY which you think are relevant. You are welcome to comment underneath.)

informative | difficult to follow | just right for me | not sufficient | crucial for my learning

helpful for my understanding | too long | unnecessarily repetitive | excessive | clear | too short | quite vague | well formulated | easy to follow | essential for my progress

Comments, if any:
……..

4) **DURING your classroom sessions, considering the time available, did you feel that the speaking activities were:** *(Please highlight ANY which you think are relevant.)*

not necessary | too short | crucial for my learning | fun | excessive | enjoyable | well-organised | essential | repetitive | useful for understanding how Russian works | boring | challenging | insufficient | helpful for my grammar | difficult to do | too long

Comments, if any:
…….
……
……..

5) **How do you think transferring to online teaching, due to COVID-19 lockdown, affected your learning within our project? Please think how the last 4 sessions that we had online, were different from our classroom sessions.** *(Mark **ONE** choice only.)*

Not affected at all: I carried on learning as before and got the most of the last 4 sessions.	
Affected **in the worst possible way**: I couldn't really carry on learning and feel I have not learned anything within the last 4 online sessions.	
Affected **mildly**: it took me a little bit to get used to the new format of learning, but I feel I have coped with it successfully and did not really miss out much during the last 4 sessions.	
Affected **quite a lot**: I felt less comfortable learning online and that I was missing some of the explanations. This meant I could not always follow the tasks and needed additional explanations. I feel that, during the last 4 sessions, I have understood some of the material but missed some other parts.	
Affected **severely**: I never got used to learning online; I was missing more information than I was following; some tasks were completed without me knowing what I was supposed to do. I feel that out of the last 4 sessions I gained very little.	

Comments, if any:
…….....
……
…….

6) **AFTER completing 20 sessions of learning Russian, please rate the following language aspects on their difficulty for you, as a learner, using the scale from 1 (*not being difficult at all*) to 10 (*being the most difficult*).** *(Rate EACH aspect in the box to the right.)*

gender agreement (e.g., мой/моя, русский/русская)		verb suffixes (e.g., работаю, работает)		reading-in-Cyrillic	
palatalization		word stress		noun suffixes (e.g., "в Москве" / "в Москву")	
noticing suffixes when reading		flexible word order when reading		prepositions (e.g., "в" / "на")	
		using suffixes when speaking			

Comments, if any:
……..
……

7) **Considering that the aim of the course was to help you learn to *read* and *speak* Russian, how do you think the course addressed the above difficulties within the time available? Please rate EACH aspect, on both explanations and practice, using the following scheme:**

1 *– did not address*
2 *– not enough time given*
3 *– appropriately*
4 *– too much time given*

	explanations	practice		explanations	practice
gender agreement (e.g., мой/моя; русский/русская)			word stress		
palatalization			using suffixes when speaking		
noticing suffixes when reading			reading-in-Cyrillic		
verb suffixes (e.g., работаю, работает)			noun suffixes (e.g., "в Москве" / "в Москву")		
flexible word order when reading			prepositions (e.g., "в" / "на")		

Comments, if any:
……..
……
…….

8) **If you could name only ONE advantage of this language course, what would you say?**
……..

9) If you could change ONE thing about this course, what would that be?

……..

THANK YOU FOR YOUR INTEREST IN LEARNING RUSSIAN AND FOR TAKING PART IN THIS PROJECT

Appendix H

EFFECTIVE WAYS OF TEACHING RUSSIAN TO ENGLISH SPEAKING BEGINNERS

QUESTIONNAIRE
Part 2

10) Your name ……

11) Did you enjoy learning a foreign language/s (e.g., French) at school?
YES/NO

Which grade/s did you get, if you remember? …… ……
(language)
(grade)

……

……

12) Now that you have completed at least 20 sessions of Russian, do you feel that in your head:
Please mark one answer (place a "V" in a box)

a) you have a lot of scattered bits of knowledge about how Russian works; ☐

b) bits of knowledge about Russian are grouped together, but some are floating in between these groups; ☐

c) you have an emerging system, not necessarily complete, but every bit of knowledge has its place; ☐

Comments, if any: .
……..
……
……..

13) When things are explained to you (by a teacher or in the handout), which of the following do you do? *(Please rate the statements below using **the full range** of the following scale:*

1 - never or almost never true of me;
2 - generally not true of me;
3 - somewhat true of me;
4 - generally true of me;
5 - always or almost always true of me.)

try to memorise the explanation and repeat it in my head; ☐

link new things to something I already know; ☐

tend *not* to ask questions; ☐

look out for patterns; ☐

take notes; ☐

compare the information about Russian to English or a foreign language I did in the past; ☐

don't always follow all the details, but do get the gist; ☐

analyse the information and check whether I understand everything; ☐

want to skip it, as I prefer to speak or read something in Russian instead; ☐

even if I understand the explanation, I might try geting additional information or further clarifications (by asking questions or by looking things up on the Internet); ☐

try to apply the new knowledge straightaway (e.g., add new suffix to the words I know); ☐

ask questions if I don't follow something; ☐

try to understand why it is like this; ☐

Comments, if any: .
……

14) Do you notice noun suffixes, when you READ in Russian independently? *(Highlight ONE please.)*

a) No, I don't ever notice them.

b) Occasionally.

c) Sometimes I do, but sometimes I don't.

d) I notice them quite a lot.

e) Yes, I would say I usually do. Perhaps I might miss one or two.

15) Please rate the statements below about what happens when you READ Russian.
Use the following scale (please try to use the full range of the scale):

1 - never or almost never true of me;
2 - generally not true of me;
3 - somewhat true of me;
4 - generally true of me;
5 - always or almost always true of me.

I struggle to remember the meanings of words, so suffixes become secondary. ☐

I get confused what the suffixes mean. ☐

I tend to analyse why the suffixes are there. ☐

I feel that suffixes in Russian are as important as words, as they add meaning. ☐

When I see an suffix, I feel I've seen it before, but I either don't recognise it or it takes me a while. ☐

I understand the rules, but I don't always think of them, when I read. ☐

I think that words are much more important than suffixes. ☐

I try to notice suffixes, but sometimes I forget. ☐

I seem to know what suffixes mean. ☐

It takes a lot of effort to read in Russian, so I have no time left to think about suffixes. ☐

I concentrate a lot on translating what I read, so suffixes tend to slip away. ☐

I tend to notice suffixes more in familiar words. ☐

I try to look out for what's at the end of words. ☐

I tend to ignore them, as I don't really understand them. ☐

I see some suffixes and know what they are, but let go of others. ☐

In Russian suffixes seem to be more important than in English, but I am not used to thinking about them, so I do not necessarily notice them. ☐

Comments, if any: .
…….
……
……..

16) When you SPEAK Russian,

a) *how often do you think about noun suffixes (e.g.,* "в Москве", "в Москву")?
(*Highlight ONE please*).

never | occasionally | sometimes | often | always try

Could you possibly give a reason for your answer?
…….
……
……...

b) *do you think about noun suffixes… (e.g.,* "в Москв**е**", "в Москв**у**")
(Delete as appropriate or put a "V" in the last box; you can only keep ONE "Yes")

less than about verb suffixes?	YES / NO
more than about verb suffixes?	YES / NO
as much as about verb suffixes?	YES / NO
I don't think I can tell. *(place a "V" in a box if this is your choice; leave blank otherwise)*	

Why do you think this is?
……
……
……..

c) *if you compare verb suffixes (e.g.,* работа**ю**, работа**ет**) *and noun suffixes* (e.g., в Москв**е**, в Москв**у**), *do you think that noun suffixes are:*
(*Highlight ONE please.)*

more difficult than verb suffixes | as difficult as verb suffixes | easier than verb suffixes | as manageable as verb suffixes

Do you think you could give a reason for your answer?
.....
……
…..

17) To choose a noun suffix, you often need to consider the following things - see below. Could you please mark them according to the order, in which you feel you make your choices?

0 *- do not think about this at all;* ***2*** *- consider this next;*
1*- decide on this first;* ***3*** *- think about it last*

gender of the noun (e.g., masculine, feminine)	
context of the sentence (e.g., location, direction)	
preposition (e.g., “в”, “на”)	
verb which is used in the sentence	

18) If you believe that you do not always use noun suffixes when you SPEAK (or do not always use them correctly), what do you think the reason/s is/are?
Use the following scale (please try to use the full range of the scale):

1 - never or almost never true of me; *4 - generally true of me;*
2 - generally not true of me; *5 - always or almost always true of me.*
3 - somewhat true of me;

I can explain which suffix I need in which context, but when I speak, they get confused in my head. ☐

I think that I don’t have to have all suffixes perfect to be able to communicate, but I do need to get at least some of them right. ☐

I think I am OK when there is one suffix in a sentence, but when there are more, I get lost. ☐

I believe that verb might help me decide which suffix the noun needs, so it’s easier with a verb ☐

I don’t think I quite understand which suffix should be used when and say them randomly. ☐

When I speak, I have an English sentence in my head and make sure that I have all the corresponding words in my Russian sentence ☐

I get suffixes correctly in the phrases that I have learned and struggle in the unfamiliar ones. ☐

We now have so many different contexts when an suffix can be used, that I feel I can't choose between them. ☐

If I make a mistake, I often know what I should have said. ☐

Before I say something, I try to think ahead to work out which suffixes I need, and occasionally get it wrong. ☐

I believe that, if I don't use the suffixes right, Russians would understand me anyway. ☐

I feel I get the first suffix ("Е" in "в Лондоне") correct more often, but haven't got used to using the second set ("У" in "в Москву") yet. ☐

By the time I remember the words that I need, it's too late for me to think about suffixes. ☐

I tend to say the first suffix which comes into my head. ☐

I miss more suffixes when I have to remember what I am supposed to say as well (e.g., reporting back on pair work). ☐

I think that Russians would struggle to understand me if I don't use the suffixes correctly. ☐

Comments. *(They would be very much appreciated)* Also, any other reasons, which I haven't thought of, are very welcome. ……..
……
……
……..
……..

19) If you believe that you do not always use noun suffixes in your own Russian speech (or do not always use them correctly), what do you think would help you improve?
(You can choose more than one reason – place a "V" in a box to the right.)

Reading more	☐	Making more sentences of my own	☐
Learning texts by heart	☐	Doing more grammar exercises	☐
Spsuffix more time on my Russian	☐	Speaking slower	☐
Learning a sample phrase	☐	Knowing the whole declension tables	☐
Practicing with someone	☐	Step-by-step guide on how to handle suffixes more effectively	☐
Developing a strategy	☐	Your own suggestion *(Please comment)*	☐

Comments, if any:
……..
……
……..

20) If you had a chance to do this course again, what would you do differently and why?

……
…….
……
…….
……
…….

21) If you have any other comments about your experience of learning to speak Russian and using noun suffixes in your speech, for example any tips or techniques you have used, or possibly some connections that you've made, or perhaps you can think of advice you would give other learners, could you please share it with us here?

……..
……
……
……
……
……

THANK YOU VERY MUCH

for your time and effort that you put into learning Russian with me, for being such great learners, for staying with me throughout these two semesters, for doing all the tests, and finally, for taking time to fill in this questionnaire. All this is greatly appreciated.

Appendix I

In-depth Interview
Participant 2029

1) Joining this project was quite a ***commitment***, what was your motivation behind it?

2) ***do you enjoy learning*** new things? What do you like most about learning? What is the hardest bit?

3) If I remember right, you did French at school and did very well in it? How do you think learning Russian compares to learning ***French***?

4) In your Questionnaire you said that you never (or almost never) "*compare the information about Russian to English or a foreign language you did in the past".* Why not?

5) I quite like your comment in your Questionnaire, "*I feel like I have a structure of Russian knowledge, with gaps to fill. And I am keen to fill them!"* Do you think you could describe that system or talk me through it?

6) You also ticked that you "***look the information up*** *elsewhere*". Could you possibly say what kind of things you looked up for Russian?

7) If we talk about reading Russian first, in your Quest. you ticked as "generally true for you" "*it takes a lot of effort* ***to read Russian***". What particularly was difficult: Cyrillic letters, pronunciation, recognising the words or figuring out grammar structures?

8) In your Learning diary/journal, one of your main concerns was pronunciation, but you ***never*** actually mentioned ***reading***. Do you think you could say why?

9) Among the things which ***could possibly help*** you with noun suffixes, you listed "*reading more*". How do you think it would help? NOTICING

10) From all the answers, related to ***noticing suffixes***, it appears that suffixes are not at the front of your mind when you read. Yet, you know that they add meaning. Could you possibly try to explain what is going on in your head when you read Russian?

11) Do you ***notice*** suffixes ***when others speak***?

12) Which of the two, **reading or speaking** did you find more difficult? Did you experience that with ***French***? Why do you think this is?

13) If we now talk about SPEAKING, do you think you could possibly ***compare*** learning to speak ***French*** and learning to speak Russian? Is one easier than the other or speaking either is difficult?

14) You are one of the top learners in the group. You put down that most of the time, you understand the rules, and yet the ***choice of noun suffixes*** can still be problematic. Have you ever thought why that is so?

    You said that noun suffixes were "***More difficult than verb suffixes***", but you were not sure why. Your comment was "***they stick less easily***". Have you had any thoughts since?

15) In our questionnaire, "5" rating meant that it is "always or almost always true for me". You only used it once to mark "***try to memorise the explanation*** *and repeat it in my head*", but then you also liked my algorithm, as well as my mneumonic "У-у-у" for "going too-oo-oo-oo". Do you prefer ***clearly formulated rules***? Or does it vary?

16) You also rated as applicable to you "*try to* ***apply the new knowledge*** *straightaway (e.g., add new suffix to the words I know)*" Could you possibly expand on that or give an example please?

17) In one of your comments, you wrote "*I like to be accurate and get things right, it is important*" and I know you get frustrated when you get the suffixes wrong. Do you have a ***strategy on how you choose*** a noun suffix?

    And perhaps on what you do if you ***make a mistake***?

    "Before I say something, I try to think ahead to work out which suffixes I need, and occasionally get it wrong."

18) In your diary, your commented that you "*felt clearer about the suffixes* ***AFTER your main speaking test***." Do you think you could probably think why?

19) When you speak, do you try to ***replicate*** the phrases that were in the text/exercise or construct your own? Do you think you try to ***translate*** from English, when you speak?

20) If I understand it correctly, you don't *believe that* ***verb might help*** *you decide which suffix the noun needs".* You said that when you choose an suffix you think of the ***verb*** last. Could you talk me through this please?

    And yet it gives you the context to differentiate, for example between, say, working somewhere and going TO somewhere.

21) The ***first comment on the difficulty of grammar*** in your diary, appears after you completed your first speaking test, when you were already using noun suffixes.
    Three consecutive entries:
    "Am getting to grips with some of the grammar, but it keeps getting more complex!"
    "Struggling a bit with the extra grammar rules like changing to y rather than e."
    "Completely lost track of what my suffixes should be and when."

    Why do you think the second set of suffixes (***zero + "Y" set***) was more challenging?

22) Do you think if the suffixes were presented in a ***table with case names***, so you could see all of them at the same time, could that have helped you use them better in your speech, or would that have made it more difficult?

23) Do you do regular work on Russian at home? What exactly do you do to speak Russian? How were you speaking?

24) Quite a number of times in your learning diary you said that "little and often and repeating sections a couple of times helps". Could you tell me more about it? What exactly do you repeat? Do you have a set number of times or does it vary?

Is this your strategy with any learning?

25) You said a couple of times that a "***step-by-step guide***" might help you improve the use of noun suffixes, but you have not ticked "developing a strategy". Why not?

Do you think they are different? Or do you think you have already got a ***strategy***?

Have you tried developing ***a strategy for noun suffixes***? Do you develop your own strategies for learning new things?

26) A lot of people find learning Russian difficult, but you said it in your diary several times and on numerous occasions in the class, that ***you enjoy it so much***. Why?

You also are ***doing so exceptionally well***. What do you think is the main reason behind it?

"I might as well use the opportunity and see how it goes rather than waste it!"

Appendix J

Code Book for Interviews

Code	Meaning	Choices
A2	enjoyed learning languages at school	Yes/No
A3	feels that they have an emerging system	a) have scattered bits of knowledge; b) some bits of knowledge are grouped together, but some are floating in between these groups; c) have an emerging system
A4	compares Russian to other languages	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A5	looks out for patterns	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A6	tries to understand why it is like this	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A6a	analyses why the suffixes are there	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A7	notices the suffixes when reading	a) No, I don't ever notice them; b) Occasionally; c) Sometimes I do, but sometimes I don't; d) I notice them quite a lot; e) Yes, I would say I usually do. Perhaps I might miss one or two.
A8	gets confused what suffixes mean (Function Assignment)	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A9	concentrates on the meaning at the expense of the form	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.

A10	thinks of the suffixes when speaking (noticing)	a) never; b) occasionally; c) sometimes; d) often; e) always try
A11b	thinks of verb suffixes more	a) less than about verb suffixes; b) more than about verb suffixes; c) as much as about verb suffixes; d)I don't think I can tell.
A11c	noun suffixes are more difficult than verb suffixes	a) more difficult than verb suffixes; b) easier than verb suffixes; c) as difficult as verb suffixes; d) as manageable as verb suffixes
A12	has a strategy for assigning a case suffix	1 - gender, 2 - preposition, 3 - context, 4 - verb
A14	no difficulties in Function Assignment	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A16	applies rules to new words	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A18	processing load	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A21	links new information to something they already know	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A21a	tries to apply the new knowledge (e.g. add new suffix to the words they know);	1 - never or almost never true of me; 2 - generally not true of me; 3 - somewhat true of me; 4 - generally true of me; 5 - always or almost always true of me.
A22	feels that they need to develop a strategy with regard to suffixes	Yes/No

Appendix K

Summary of the results of the Part 2 of the Qualitative Questionnaire (the codes are listed vertically and participants – horizontally).

	2001	2002	2004	2006	2007	2008	2009	2012	2015	2017	2018	2019	2020	2021	2023	2024	2025	2026	2027	2028	2029
A2	Y	N	Y	y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	Y	Y	Y
A3	C	b	c	B	B	C	C	B	A	C	c	B	C	C	B	B	B	B	C	B	C
A4	5	4	5	2	4	4	5	5	4	1	3	5	1	4	5	5	3	3	5	1	1
A5	2	4	5	5	4	5	5	5	3	4	5	4	4	3	5	4	3	2	4	5	4
A6	2	4	5	3	3	5	5	5	5	4	2	2	3	5	3	3	2	4	4	5	3
A6a	1	4	3	3	3	4	4	5	3	1	2	2	2	4	3	4	3	3	4	4	3
A7	e	E	c	D	C	e	E	d	D	D	D	B	D	D	E	B	C	C	E	E	E
A8	1	2	2	3	3	1	2	3	3	3	2	2	2	2	2	2	3	3	3	2	3
A9	3	3	5	2	3	3	3	4	2	3	4	3	4	3	3	3	2	2	2	3	2
A10	e	e	e	D	d	e	E	e	e	D	E	B	E	E	E	E	E	D	D	E	E
A11b	a	a	c	a	b	A	A	B	A	B	A	A	C	D	A	D	C	A	B	A	C
A11c	a	d	a	B	a	A	A	A	A	B	D	A	C	D	D	A	A	C	A	B	A
A12 gender1 preposit2 context3 verb 4	1, 2, 3, 0	4, 1, 3, 2	4, 3 2, 1	1, 3, 4, 2	4, 3 2, 1	4, 3, 1, 2	3, 1, 2, 4	3, 1, 4, 2	4, 3, 2, 1	1, 2, 3, 4	1, 2 3, 4	4, 2, 3, 1	4, 2, 3, 1	3, 2, 1, 4	2, 1, 3, 4	4, 1, 3, 2	2, 1, 4, 3	4, 1, 3, 2	3, 4, 2, 1	4, 1, 3, 2	1, 3, 2, 4
A14	4	4	4	4	4	4	4	4	3	4	3	3	4	4	4	3	3	3	5	4	4
A16	5	3	2	4	3	3	3	3	2	3	3	3	2	5	5	2	3	3	5	2	4
A18	4	4	5	5	3	1	3	5	3	2	4	3	4	2	3	3	2	3	2	2	3
A21	5	4	5	5	4	4	4	5	4	2	4	3	1	4	5	3	4	3	4	4	4
A21a	5	3	2	4	3	3	3	3	2	3	3	3	2	5	5	2	3	3	3	2	4
A22	Y	N	N	N	Y	N	N	Y	Y	N	N	Y	N	Y	Y	N	N	N	N	N	N
WM span	38	49	50	50	48	68	44	69	33	23	48	15	48	48	48	61	65	65	69	62	68
Test 2 rate	68	89	87	93	88	95	90	74	80	69	60	52	88	71	84	83	94	67	86	68	90

Appendix L

Russian Case Inflection Teaching Syllabus
for Ab initio (Year 1) course,
following the principles of the proposed Spiralling Teaching Framework

Proposed Spiralling Syllabus

Ab initio (Year 1)

This syllabus is created following the principles of my proposed Spiralling Grammar Teaching Framework (see Section 4.2).

As, in many UK Universities, students starting a beginners' Russian course are expected to be able to read Russian, this proposed syllabus does not include teaching the Cyrillic alphabet. However, the hours for teaching to read Cyrillic can be accommodated at the expense of Week 21 and Week 22, when no new case forms or verb forms are taught and the aim of the teaching is to further develop the four language skills, as well as expanding students' active vocabulary and improving the language use through mainly communicative activities. The introduction of the Cyrillic alphabet and Russian pronunciation, parallel to basic vocabulary and very basic grammar, using spiralling approach, is presented in my textbook "Russian in Plain English: a Very Basic Russian Starter for Complete Beginners (2020) by Routledge.

Week	**session**	**Grammar**		**Topic**
		Gender & Case	**Verb / Syntactic structures**	
T E R M 1				
Week 1	1	Gender (masculine, feminine)	No "to be" in Present	World around us: things and people
	2	Possessive pronouns "my" / "your/s" and their agreement with the two genders	Introductory sentences "This is"	World around us: places
	3	Word order, questions, Nominative, Yes/No answers	"To eat" (1st & 3rd person Singular) "â em", on/a est"	World around us: food items
	4	Question words "Who?", "What?", "Where?" with "tam" (for "*there*") and "doma" (for "*at home*") as answers	"I/you have" constructions	World around us: basic shopping conversations
	5	**Prepositional case:** "**e**" suffix with masculine and feminine nouns, e.g.,"v Moskv**e**" (for "*in Moscow*"); Preposition "v" for "*in*";	-	World around us: people and places

		Monosyllabic “v sad**u**” (for “*in the garden*”)		
Week 2	6	Prepositional case: preposition *na* for “on”; Opposition of Prep. and Acc in masculine (musical instruments vs games)	Verb conjugation (1st): Infinitive; 1st person Singular	What we do: sports games; musical instruments
	7	Possessive pronouns “his” /”her/s” (no agreement with the noun is required); prepositions *v* and *na* for “at”	Verb conjugation (1st): 3rd person Singular	What we do: work and leisure; jobs
	8	Adjectives and their gender agreement in Nominative; neuter nouns; 3rd declension feminines (with the Soft Sign)	Monosyllabic verb “žit’”; “est’” as “*there is/are*”	What we do: places where we live
	9	Plural – “y” / “i" suffixes; 7-letter spelling rule	Verb conjugation (1st): 2nd person Singular 2nd person Plural	What we do: turn-taking and conversations
	10	“He/she has” construction; adverb vs adjective	Infinitive after *lûblû* (for “I like”)	What we do: hobbies
SPEAKING TEST 1 (see 5.4.3) on Prepositional suffix “e” as opposed to Nominative (base form) (can be recorded and emailed to the instructor; can be formative or summative)				
Week 3	11	**Accusative case** (scaffolding for masculine – zero-suffix) with days of the week; preposition *v* with days of the week; conjunction *potomu čto* (for “because”)	2nd conjugation: “to like” in 2nd and 3rd person Singular	Our schedules: days of the week
	12	Accusative with feminine days of the week - overt feminine suffix “**u**”	2nd conjugation: “to speak” and “to learn (a language)”	Our schedules: working week, languages
	13	Accusative in directionality context (all three genders), e.g. *idët v universitet_* (for “is going to University”); *Idët v školu* (for “is going to school”)	Verb *idti* for “to go” (on foot or within a walking distance) – all Singular forms + 2nd person Plural	Our schedules: going places
	14	Accusative (directionality) vs Prepositional (locationality); Prepositional with transport	Verb *ehat'* for “to go” (by transport) – all Singular forms + 2nd person Plural	Our schedules: means of transport

	15	"Where?" in directionality and locationality contexts (*Kuda?* vs *Gde?*); Prepositional with months; Accusative in the object function (masculine and neuter – zero suffix) – "to watch TV", "to listen to the radio"	2nd conjugation (not standard): Verbs *smotret'* (for "to watch (TV)"), *sidet'* (for "to sit"), ležat' (for "to lie")	Our schedules: time off and travelling; months
Week 4	16	Numbers and times	Verbs "to get up", "to sleep"	Every day: basics
	17	Accusative feminine in the object function	Verbs "to eat" and "to drink" (in Singular and in 2nd person Plural)	Every day: meals
	18	Nominative vs Accusative in the object function (Subject – Object); word order	1st conjugation: 1st person Plural; Verbs for having meals, e.g., *zavtrakat'* (for "to have breakfast")	Every day: doing things together
	19	Irregular Plurals with "**a**"/ "**â**" suffixes in Accusative, e.g., *âjc**a*** (for "eggs"), *hlop'**â*** (for "flakes"); frequency adverbs, e.g. "usually", "often", "sometimes"; Accusative in "every day" (masculine)	2nd conjugation: 1st person Plural, incl. "we eat" (irregular);	Every day: going to a shop/ restaurant/cafe
	20	"We/they have" construction; Possessive pronouns "our/s" and "their/s" and their agreement	3rd person Plural – 1st and 2nd conjugation	Every day: meals in Russia and in Britain
Week 5	21	Prepositional with "iâ" feminines (subclass with a different suffix- "ii"); names of countries; Prepositional Plural, e.g. "on holidays" ("**ah**" suffix)	Past Tense: *byl/á/o/i (*for "was/were")	Holidays: basics
	22	Prepositional in time references for the past, e.g. last week, last year (including the adjective *prošlyj* (for "last")); clauses with "when"	Standard verbs in Past Tense	Holidays: activities
	23	Accusative for "iâ" subclass ("**û**" suffix – allophone of "u"); Prepositional vs Accusative with places (*byl* (for	Verb "to go" in the Past –"*hodil/a/i* and *ezdil/a/i* (for "went")	Holidays: travelling

		"was/were") vs *hodil*/a/i and *ezdil/a/i* (for "went"))		
	SPEAKING TEST 2 (see 5.4.4) Prepositional vs Accusative (suffixes and contexts studied by this point) (can be formative or summative)			
	24	**Instrumental case**: Singular (nouns of all three genders); preposition *s* for "with", e.g., *s drug**om*** ("with a friend (male)"), *s podrug**oj*** (for "with a friend (female)"	Past vs Present; The verb *otdyhat'* (for "to have a break", "to have a holiday", "to relax")	Holidays: my favourite trip
	25	Instrumental case: Plural, e.g., *s druz'**âmi*** (for "with friends"); Questions using "With whom?" (Instrumental of "who?") Prepositional + Instrumental	The verb "to want" (Singular)	Holidays: the best day in my life
Week 6	Reading week			
Week 7	26	noun+noun phrases in Russian, e.g., *buterbrod s syrom* for "cheese sandwich" (requiring Instrumental); Instrumental: Singular vs Plural (uncounTable 12s countable); Personal pronouns in Instrumental *s nim* (for "with him"), *s nei* (for "with her")	The verb "to cook"	Food & Cooking: basics
	27	Instrumental: 3rd declension (Soft Sign feminine) – "***'û***" suffix, e.g., *s sol'**û*** (for "with salt")	The verb "to want" (Plural)	Food & Cooking: Russian cuisine
	28	Accusative + Instrumental; Questions using "With what?" (Instrumental of "what?")	The verb "to give"; Imperative; Subjectless structures *možno/nel'zâ* + Infinitive"(for "It is (not) possible/allowed")	Food & Cooking: ordering food
	29	**Accusative** in frequency references, including an **adjectives**, e.g., *každ**uû** nedel**û*** (for "every week"); Prepositional for 3rd declension (Soft Sign feminine), e.g., *v očered**i*** (for "in the queue")	Verbs "to buy" (1st conjugation) and "to pay" (2nd conjugation)	Food & Cooking: buying food

	30	**PRESENTATIONS** (5 minutes + 2 questions) on “Food” topics, with focus on Instrumental but also testing the use for Prepositional and Accusative (normally in-class and graded; but could be formative)		
Week 8	31	Ordinal numerals (adjectival suffixes) and their agreement, e.g., “first semester”, “first lecture”, “second seminar”; subjects	“have”-construction in the Past Tense and their gender agreement	My life at University: timetable and curriculum
	32	**Genitive Singular** in “of-constructions” with masculine – “**a**”/ “**â**” suffixes; Dates (months are masculine); Genitive with negation (none of), e.g., *U nas ne bylo uroka/signal* (for “We did not have a lesson/a signal)	Negative forms of “have”-construction (requiring Genitive);	My life at University: pros and cons
	33	Genitive feminine Singular – “**y**”/“**i**” suffixes; 7-letter rule; Genitive with containers and amounts, e.g., *stakan soka* (for “a glass of juice”), *butylka vody* (for “a bottle of water”)	-	My life at University: my shopping basket
	34	Accusative + Genitive, e.g., *pokupaû butylku vody* (for “I buy a bottle of water”); Prepositional with “pokupat’”, e.g., *pokupat’ v magazine* (for “to buy from a shop”)	Im/Perfective Aspect: *pokupaû* (for “I buy”) vs *kupil* (for “I [have] bought”)	My life at University: places to shop
	35	“ne/daleko ot” (for “*(not)far from*”) + Genitive; Instrumental (preposition “with”) vs Genitive (preposition “without”)	*zabyl/nado kupit’* (for “forgot/need to buy”)	My life at University: my fridge
Week 9	36	The preposition *o* “about” + Prepositional; Genitive with authors, e.g., *roman Pushkina* (for “novel by Pushkin”); “mâ”-neuter, e.g. *svobodnoe vremâ* (for “free time”)	Im/Perfective – prefixed verbs	Hobbies - Books
	37	Instrumental for tools, e.g., *risovat’ karandašom/ /kraskami* (for “to draw with a pencil/to paint with paints”)	“-ova”-verbs, e.g., *risovat’* (for “to draw, to paint”)	Hobbies - Arts and Music

	38	Instrumental with "*zanimat'sâ*" e.g., *zanimat'sâ sport**om**/muzyk**oj*** (for "to do sports/music"); adjectives in *zanimat'sâ russk**im** âzykom* (for "to study Russian language")	The reflexive verb *zanimat'sâ* (for "to do", "to be engaged in", to occupy yourself with")	Hobbies - Sports
	39	Accusative with events - *na* vs *v,* e.g., *hožu v teatr/na koncert* (for "go to the theatre/to a concert")	Multidirectional verbs of motion	Hobbies - Going out
	40	**Traditional Written GRMMAR TEST 1** on the case forms studied (full set of nominal suffixes for Prep., Acc. inanimate, Instrumental and Genitive Singular, including 7-letter spelling rule)		
Week 10	41	**Accusative Animate** (masculine and feminine Singular)	The verb *vstrečat', vstretit'* (for "to meet")	Having fun: my friends
	42	Evaluations of feelings and emotions, e.g., *budet veselo* (for "that will be fun")	Future Tense (Imperfective)	Having fun: plans for the evening
	43	Acc. adjectival in directionality context	Verbs of Motion in Future – *pojdu* and *poedu* (for "I'll go")	Having fun: plans for the week
	44	"mnogo" (for "*a lot of*") + Genitive with uncountables, e.g., *budet mnogo piv**a*** (for "there will be lots of beer")	"I have "-construction in the Future Tense, e.g., *U menâ budut gosti* (for "I will have guests/visitors")	Having fun: planning a party
	45	Numerals 2, 3, 4 with Genitive Singular (scaffolded)	Future Perfective	Having fun: planning to cook
Week 11	46	Preposition *na* with events and special occasions in Accusative, e.g., *na Rozdestvo* (for "for Christmas")	Future vs Present (Imperfective)	Special occasions: basics
	47	Preposition *dlâ* (for "for") + Genitive; Instrumental (preposition *pered* for "before") vs Genitive (preposition *posle* for "after")	Future vs Past (Perfective), e.g., *kupil/a/i* (for "bought"), *kuplû* (for "will buy")	Special occasions: Christmas shopping, presents
	48	Instrumental after *pozdravlât'*, e.g., *pozdravlâu s Novym*	The verb *pozdravlât'* (for "congratulate")	Special occasions:

		godom (for “Happy New Year”)		Christmas and New Year traditions
	49	Neuter nouns with “’e” suffix, e.g., *zdorov’e* (for “health”); Genitive “**â**” suffix for neuter	The verb “to wish” + Genitive	Special occasions: Christmas/New Year cards
	50	**Loosely structured Interactive Test** conversation with a peer “Plans for Holidays”, 8-10 minutes, tests all the case forms which were tested in the last written Grammar Test + animate Acc. and Future Tense (can be formative or summative)		
CHRISTMAS BREAK (End of Semester 1) The first essay 150 words				
Week 12	51	Vocabulary for furniture	The use of verbs “to stand”, “to lie” and “to hang” instead of “*there is/are*”	Where we live: my room
	52	**Adjectives and ordinal numerals in Prepositional** (masculine), e.g. *na pervom ètaže* (for “on the ground floor”)	The verb *nahodit'sâ* (for “to be situated”)	Where we live: our hostel
	53	The agreement of the pronoun *eto* (for “this”) with nouns; Adjectives and pronouns in Prepositional (feminine), e.g., *v ètoj komnate (*for “in this room”)	-	Where we live: flats in Russia
	54	Three prepositions for “from” + Genitive Adjectival with Accusative (masculine), e.g., *pereehali v novyj dom* (for “moved into a new house”	Verb *pereehat’* (for “to move (house)”)	Where we live: moving house
	55	*možno nomer na* + Accusative?” (for “Can we have a room for..?”)	The verb *priehal/a/i* (for “arrived”)”	Hotel rooms
Week 13	56	Numerals from 100 to 1000; Years in Russian; Ordinal numerals in Prepositional, e.g., *v 2020-**om** god**u*** (for “in the year 2020)	The verb “to be born” (reflexive in Past Tense)	Biographies – year of birth and birthplace
	57	No Accusative with reflexive verbs;	Reflexive verbs: *učit'sâ, zanimat'sâ* for “to study”	Biographies – education

		Prepositional with “study” verbs;		
	58	Genitive in possessive phrases with animates, e.g., “friend’s sister”	Verbs for “getting married” for males and females	Biographies: family tree
	59	**Dative** (nouns) with the preposition *k* (for “to/towards” (masculine, feminine and Plural), e.g., *ezdit’ k brat**u**/ k sestr**e**/ k roditelâm* (for “to go to see brother/sister/parents”)	-	Visiting family & friends
	60	Mini-presentations (normally formative)	All three tenses must be used	Our relatives and friends and where they live
Week 14	61	Dative with no preposition	Verbs using Dative, e.g., “to ring”, “to give”, “to say”	Communication: landline and mobile
	62	Dative vs Instrumental, e.g., *skazal/a brat**u**[Dat]*” (for “told my brother”) vs *govorila s brat**om**[Inst]* (for “spoke with my brother”); Personal pronouns in Dative	speech verbs	Communication: telling stories
	63	Dative vs Genitive, e.g., *pis’mo brat**u**[Dat]* (for “the letter to my brother”) vs *pis’mo ot brat**a**[Gen]* (for “the letter from my brother”)	The verbs “to receive”, “to send” (Imperfective/ Perfecive)	Communication: our correspondence
	64	Preposition *po* (for “on” as “via”) + Dative, e.g., *po telefon**u*** (for “on the phone”	The verbs *zvonit’* and *zvat’* for “to call”	Modern technologies
	65	**Speaking Test 3** a simple video episode about visiting friends/family can be summarized in 7-10 sentences, tests Dative, some adjectival case suffixes studied and the three tenses (can be formative or summative)		
Week 15	66	Parts of the body vocabulary; Adjective agreement in Plural Nominative; The question word *kakoj* (for “which? what kind of?”)	“I have”-structures with no *est’*	Body and mind: how do we look?
	67	Genitive in “having” structures, e.g. *u sestr**y**[Gen]* –		Body and mind: how do we look?

		krasivye volosy (for "My sister has beautiful hair", Plural)		
	68	Stress shift in Plural; more Irregular Plurals;	The verb *bolet'* (for "to be ill" and "to hurt/ache") – two ways of conjugation	Body and mind: aches and pains
	69	Dative for describing a state, e.g., *mne ploho* for "I feel ill"; Personal pronouns in Dative.	"Change-of-state" Perfective forms, e.g., *zabolel* (for "fell ill"), *uvidel* (for "suddenly saw")	Body and mind: visit to a doctor
	70	**Adjectives in Dative**, e.g., *nado idti k zubn**omu** vrač**u*** for "I need to go to a detist"	Future Perfective of "give", e.g., *Vrach dast mne lekarstvo* (for "the doctor will give me a medicine")	Body and mind: visit to a dentist
Week 16	71	**Genitive Plural** masc. "ov" and feminine & neuter zero suffix (drops the base-form marker "a"/"o") with *mnogo* (for "much/many/lots of") or "net" (for "there is no")	"There is/are" construction in the Past	Urban vs countryside: differences
	72	Vowel insertions in Genitive fem. Plural, e.g., *mnogo sum**o**k_* (for "many bags"); 5-letter rule; Genitive Plural vs Instrumental Plural - *daleko ot dom**ov*** (for "far from houses") vs *râdom s dom**ami*** (for "next to houses")	The verb *stroit'* (for "to build"); Passivised "they-construction", e.g., *stroât školu* (for "they are building a school")	Urban vs countryside: pros and cons
	73	Countables and uncountables in Genitive	The verb "to sell"; Passivized *v magazine prodaût* (for "In the shop, they sell ..")	Urban vs countryside: shopping, incl. clothes
	74	Genitive Plural clusters taking the "ej" ending - a) words ending in the Soft Sign; b) masc. ending in "hushers"; c) irregular Plurals, e.g., *mnogo lûdej* (for "many people")	-	Urban vs countryside: entertainment
	75	Discussion Dative in *mne nravitsâ* construction (for "I like")	Syntax of *mne nravitsâ* construction (for "I like");	Urban vs countryside: where I would like to live and why

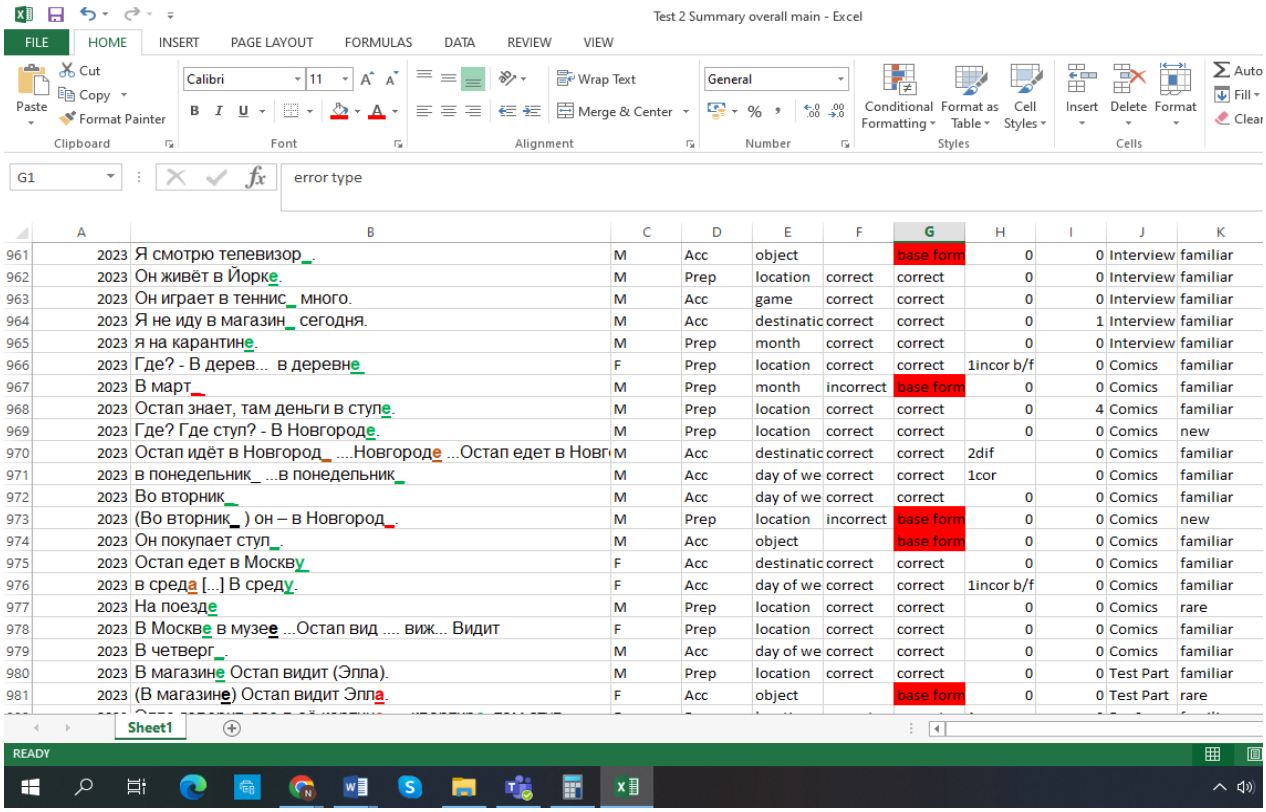
			The phrase *hotel/a/i by* (for “would like”)	
Week 17	Reading Week Preparation for a presentation using all cases and tenses			
Week 18	76	In-class presentations about different cities in Russian-speaking countries. should include unrehearsed questions from class and teacher (Genitive assessment)		
	77	Dative vs Accusative; Preposition *po* (in the meaning of “along”) vs *čerez* (for “across”)	Imperative	Navigating a city: directions
	78	Ordinal numerals for bus numbers **Adjectival Prepositional vs adjectival Accusative;**	The verb *sadit’sâ* (in the meaning of “to get on/ to board”) and *vyhodit’* (in the meaning of “to get off”	Navigating a city: using transport
	79	Genitive Singular and Plural after numerals	Im/Perfective with Imperative; the verbs “to pay” and “to spend” (money)”	Navigating a city: money and paying
	80	Work in mini-groups – planning a guided tour for Russian-speaking visitors	The verb “to take”, e.g., *voz’mi/te* (for “you can have (take) this” Future Perfective of new verbs	Navigating a city: guided tours
Week 19	81	Use of nouns to describe weather; Genitive with negation, e.g., *net doždâ* (for “it is not raining” = “no rain”)	No-verb sentences, e.g., *Solnce* (for “It is sunny.”)	Weather: basics
	82	Adjectives vs adverbs; Plural of “mâ” neuters – *vremena goda* (for “seasons”)	Subjectless sentences, e.g., Teplo. (for “It is warm.”)	Weather: seasons and outdoor activities
	83	Comparative forms of adjectives and adverbs; “snowman” as animate; Using new “winter” vocabulary in describing pictures	The phrases for “skiing” and “skating”	Weather: Russian winter; clothing
	84	**Adjectives in Genitive Plural** (all genders), e.g., *net sil'nyh morozov* (for “no severe frosts”)	-	Weather: climate in different Russian-speaking countries

	85	Pronoun *svoj* (reflexive possessive); Interactive activities	Past and Future of Subjectless sentences	Weather: my favourite season/climate and why
Week 20	86	**Adjectives in Genitive Singular** (masculine and neuter) e.g., *mnogo/nichego interesn**ogo***; Vocabulary for events, e.g., *spectakl'* (for *"*a performance"), *festival'* (for "a festival")	Imperfective Verbs with "-yva-"/"-iva-", e.g., *zakazyvat', zakazat'* for "to book, reserve"	Events: programmes and bookings
	87	Dates in Genitive (adjectival)	The verbs for "to open" and "to shut"	Events: dates and times
	88	Adjectives in Genitive Singular feminine, e.g.*, net moej lûbimoj aktrisy* (for "my favourite actress is not on")	The verb "to be late";	Events: on the day
	89	Personal pronouns in Perpositional, e.g., *v nëm* (for in him/it), *v nej* (for "in her/it), *v nih* (for "in them")	Past Imperfective vs Past Perfective	Biographies of famous composers/musicians/ writers/actors/etc
	90	"If" and "when" clauses, e.g. *kogda â poedu v London* (for "When I go to London")	Reflexives which can drop their "sâ", e.g., *načat'/câ* (for "to begin, to start")	Future events
Week 21	91	Traditional Written Grammar Test 2 on Dative, Genitive and verb forms		
	92	Relative pronoun *kotoryj* (for "which", "that") in Nominative – adjectival-type gender agreement	Subject-Verb agreements in clauses	Multicultural communication: explaining who is who
	93	Relative pronoun *kotoryj* – case agreement: Prepositional, Instrumental and Genitive (with prepositions)	Expanding verb range from communicative contexts; monitoring Im/Perfective use	Multicultural communication: explaining national cuisine and food etiquette
	94	Relative pronoun *kotoryj* – case agreement: Nominative vs Accusative; Dative and Instrumental without prepositions	Expanding verb range from communicative contexts; monitoring Im/Perfective use	Multicultural communication: asking and explaining traditions

	95	Adjectives in Genitive Plural vs Prepositional Plural; Work in mini-groups – search Russian-speaking Internet	Expanding verb range from communicative contexts; monitoring Im/Perfective use	Different Russian-speaking countries
Week 22	96	Reading and retelling pieces of Russian-speaking literature (work in mini-groups)		
	97	Listening to and retelling stories		
	The second essay on any of the topics studied tests all case forms studied in Semester 2 – Dative, Genitive and all adjectival			
	98	Interactive Activities on Oral Exam topics		
	99	Grammar Exam Preparation		
	100	Oral Exam Preparation		

Appendix M

Image of a section from Test 2 summarizing spreadsheet.



Test 2 Summary overall main - Excel

	A	B	C	D	E	F	G	H	I	J	K
961	2023	Я смотрю телевизор_.	M	Acc	object		base form	0	0	Interview	familiar
962	2023	Он живёт в Йорке.	M	Prep	location	correct	correct	0	0	Interview	familiar
963	2023	Он играет в теннис_ много.	M	Acc	game	correct	correct	0	0	Interview	familiar
964	2023	Я не иду в магазин_ сегодня.	M	Acc	destinatic	correct	correct	0	1	Interview	familiar
965	2023	я на карантине.	M	Prep	month	correct	correct	0	0	Interview	familiar
966	2023	Где? - В дерев... в деревне_	F	Prep	location	correct	correct	1incor b/f	0	Comics	familiar
967	2023	В март_	M	Prep	month	incorrect	base form	0	0	Comics	familiar
968	2023	Остап знает, там деньги в стуле.	M	Prep	location	correct	correct	0	4	Comics	familiar
969	2023	Где? Где стул? - В Новгороде.	M	Prep	location	correct	correct	0	0	Comics	new
970	2023	Остап идёт в Новгород_ ....Новгороде ...Остап едет в Новг	M	Acc	destinatic	correct	correct	2dif	0	Comics	familiar
971	2023	в понедельник_ ...в понедельник_	M	Acc	day of we	correct	correct	1cor	0	Comics	familiar
972	2023	Во вторник_	M	Acc	day of we	correct	correct	0	0	Comics	familiar
973	2023	(Во вторник_) он – в Новгород_.	M	Prep	location	incorrect	base form	0	0	Comics	new
974	2023	Он покупает стул_.	M	Acc	object		base form	0	0	Comics	familiar
975	2023	Остап едет в Москву_	F	Acc	destinatic	correct	correct	0	0	Comics	familiar
976	2023	в среда [...] В среду.	F	Acc	day of we	correct	correct	1incor b/f	0	Comics	familiar
977	2023	На поезде	M	Prep	location	correct	correct	0	0	Comics	rare
978	2023	В Москве в музее ...Остап вид .... виж... Видит	F	Prep	location	correct	correct	0	0	Comics	familiar
979	2023	В четверг_.	M	Acc	day of we	correct	correct	0	0	Comics	familiar
980	2023	В магазине Остап видит (Элла).	M	Prep	location	correct	correct	0	0	Test Part	familiar
981	2023	(В магазине) Остап видит Элла.	F	Acc	object		base form	0	0	Test Part	rare