Varieties of Initial Learners in Language Acquisition (VILLA): Controlled classroom input and elementary forms of linguistic organisation

This project is submitted in response to the call for proposals OPEN RESEARCH AREA IN EUROPE FOR THE SOCIAL SCIENCES to the funding agencies ANR, DFG and NWO. Research will be conducted in four countries (France, Germany, The Netherlands, UK) but funding is only needed for the sub-projects in France, Germany, and the Netherlands.

1. The aims/objectives of the research proposed

In a world of increasing communication and migration across countries and linguistic territorial lines a growing number of speakers have to use a second language on an everyday basis and a lot hinges on their success. If there is one linguistic problem of both exceptional scientific interest and eminent social significance, it is the problem of second language learning and teaching. There is therefore a heightened urgency to understand both the processes underlying language learning and the influence of specific didactic approaches on these learning processes.

Work at the interface between research in language acquisition and pedagogy is essential for the advancement of our practices of language teaching, enhancing language curricula, and making informed decisions about language policies affecting foreign language and multilingual education in our schools. In this way, the current project proposes to contribute concretely to the European Council's Strategic Framework for European Cooperation in Education and Training ("ET 2020") with its special focus on the promotion of language teaching.

Second language learning can only take place, when learners have access to target language speech produced by target language speakers around them. In a learning situation, this incoming stream of language is commonly referred to as 'input' (see Carroll 1999, 2001 for a problematisation of this notion and Gass 1997, Hatch 1983, Krashen 1985 for some relevant studies). We all know the experience of being exposed to a wall of unintelligible noise, e.g. when we have inadvertently switched on the 'wrong' radio station. But how are we able to get from here to a stage at which we can distinguish individual words, know what they mean, and how they can be combined to form more complex expressions?

This project sets out to investigate the very first phases of foreign language acquisition under controlled input conditions. Total beginners (adult learners with four different native languages: Dutch, English, French, German) will be exposed to 20 hours of input in a new language (Polish). The language input will be monolingual and structured in such a way as to allow the testing of students at different linguistic levels, it will be held constant under relatively natural and interactive exposure conditions, and it will be recorded and transcribed for detailed analysis.

The learners' performance will be investigated over time at different levels of language (perception, comprehension, grammatical analysis and production) in order to find out more about the most elementary stages of linguistic organisation resulting from the treatment of the input. Empirical investigations will cover phonology, lexicon, morpho-syntax and pragmatics of the developing initial learner varieties.

The project brings together a core group of researchers who have a successful history of collaboration in the context of the “Structure of Learner Varieties” research group, supported by the European Science Foundation and the Max Planck Institute of Psycholinguistics in
Nijmegen, NL (cf. www.learner-varieties.eu). This core group has reached out to colleagues with diverse expertise in language acquisition and teaching to create a new research group with a focus on the interface between these two domains, while also valorising the findings of prior collaborative work (cf. Perdue 1993, Dimroth & Starren 2003, Hendriks 2005). This research group provides the basis for the current project.

The decision to focus on the beginning of the language acquisition process has come about for numerous reasons. Firstly, very little is known about the early stages of adult language acquisition. Theories of second language acquisition lack an accurate description of the initial state and this has consequences for application. The CEFR’s level A1, for example, sets the threshold of language proficiency scaling too high. Much learning takes place before this stage, and it is this learning that interests us here. Documenting the initial stages is therefore necessary to fill this gap as suggested by numerous researchers (Perdue 1990, Vainikka & Young-Scholten 1998). Secondly, previous research in this domain has, for the most part, used artificial languages as the object of investigation; the authors of the current proposal are amongst some of the few researchers who have begun to work on the early stages of natural language acquisition, and a comprehensive cross-linguistic study of this kind has not yet been conducted. Thirdly, the influence of the learners’ linguistic environment on the acquisition process has been neglected in past research. Observing and analyzing acquisition processes from the first moments of exposure to the new language allows for a meticulous examination of the linguistic input to which learners are exposed (in the language classroom in our case) and of learners’ performance in the target language relative to this input. This can be done precisely because, with this approach, the learners’ previous target language knowledge can be kept constant (i.e. at zero).

This project therefore proposes a complete analysis of the natural target language input in the classroom, including learners’ productions during language instruction. Learners’ performance on target language tasks will be analyzed with respect to the input and to the work each learner does on this input. Extensive psycholinguistic experiments examining learners' perception and comprehension of the new language will be administered to the learners after classroom exposure at regular time intervals throughout the duration of the experimental period.

Prior linguistic knowledge has always been considered one of the core factors influencing second language acquisition processes and the nature of intermediate learner grammars. This holds in particular for the learners' native language(s). The cross-linguistic nature of the project will allow us to observe how learners with different native languages work on the same target language input. The four-country collaboration will provide us with the unique opportunity to collect data from native speakers of French, Dutch, German and English using identical tools and methodologies. This will be achieved by exposing four adult learner groups with the four different native languages to 20 hours of controlled input in Polish (The "Basic Input" groups, see Table 1.1).

In addition to this basic comparison focusing on the influence of the learners' L1 at the initial stages of L2 acquisition, two further factors will be investigated. For reasons of synergy the experiments targeting these additional factors will be distributed among the participating teams.

The effects of different approaches to presenting the same input will be studied in two of the collaborating countries (France and The Netherlands) by comparing the results from the "Basic Input" study to the results achieved by a group of adult learners under different exposure conditions. Whereas the "Basic Input" group will be exposed to structured monolingual input in which no meta-information is provided, the learners in the "Explicit
Input" group (see Table 1.1) will in addition get access to more form-related information on the structure of the new language. Their attention will for example be explicitly directed towards different morphological forms and the conditions determining their use (e.g. singular vs. plural NPs). Comparisons will be made between learners in the two groups by observing their productions in the classroom during their immediate processing of the input and their performance on the same series of tests in the target language. Different types of exposure may help the learner with different types of analyses; i.e. some types of exposure may enhance the learners’ level of accuracy, whereas others may enhance the learners’ level of fluency. In sum, it may very well be that different types of exposure to the TL will result in different types of benefits to the learner.

Another core factor of major theoretical but also applied relevance in the field of second language acquisition is the role of the learners' age. Under the assumption that rapid native-like attainment was nearly guaranteed for children, child L2 acquisition has been systematically neglected during the past decades of second language acquisition research. Recent research (e.g. Ahrenholz 2006, Philp et al. 2008, Unsworth 2005) clearly indicates that child L2 acquisition does by no means always progress without difficulty. Given that L2 learning children constitute a population of core interest for pedagogical decision making, the teams in the other two collaborating countries (Germany and the UK) will investigate language acquisition in a group of child learners (age 10) exposed to the same input as the "Basic Input" group (see "Child learners" in Table 1.1). Comparisons will be made between child and adult learners’ productions during and after classroom instruction and their performance on an identical set of target language tests.

Table 1.1. Source languages, types of exposure and age groups

<table>
<thead>
<tr>
<th>Source language</th>
<th>Basic input</th>
<th>Explicit Input</th>
<th>Child learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>English</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>French</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>German</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

The principal applied objectives of the project are to contribute to reflections on how language teachers can optimize classroom time and enhance their students’ experience of appropriating a new TL, and on how to best assess their students’ language proficiency and development. Our mission is to help train future language teachers, and in order to do this we need a clear picture of what type of input is most effective and a language instruction program that makes use of this input. We propose to begin bridging the gap between language acquisition and pedagogy by attempting to better understand the processes underlying second language acquisition in the classroom in order to work out, in a second step and in collaboration with experts in language pedagogy, how these processes can best be enhanced by language teaching.

2. The intellectual, scientific and policy background to the proposal and why it is important

The precise properties of target language input and their influence on the process of second language learning have been neglected in L2 acquisition studies for decades. The role of
input has been viewed as negligible because of a strong generative-theory based hypothesis that the L2 initial state is the equivalent of an individual’s native language (L1) final state. In other words, a learner’s L1 is the basis for learning a new L2, and input does nothing more than trigger the language learning mechanism to reset parameters, that is, to shift from the original L1 structure to the new L2 structure (cf. Schwartz & Sprouse 1996, White 2003).

This view has recently been challenged by researchers proposing a usage-based and form-focused approach to language acquisition according to which language learners are strongly influenced by the statistical distribution of surface properties in the target language input (cf. Ellis 2008, Tomasello 2003). Accordingly, the description of input properties has received a lot of attention by proponents of this line of research.

Research conducted within a functionalist framework has shown, however, that even the most basic user of a foreign language produces utterances that cannot be explained by a direct appeal to the grammar of the first language nor to that of the language being learned or used (cf. Klein & Perdue 1997, Master et al. 1989, Starren, 2006). According to the so-called 'learner varieties approach', second language acquisition always involves processes of creating language anew (Perdue 2006). Such processes are assumed to be informed by principles that are specific to language and communication, involving language neutral instead of only source- or target language knowledge: "To the extent that learner varieties also exhibit properties which are independent of source and target language, we must assume they immediately reflect creative processes of the underlying human language faculty. Where else should these properties come from?" (Klein 2001: 90).

To inform this debate, a more precise description of the first and earliest stages of L2 acquisition is necessary. This can only be done by means of controlled experiments on natural linguistic input to which learners are exposed and on their performance during language tasks relative to this input and to their L1. In his article “Give input a chance!” the influential phonologist, J.E. Flege (2009: 190), concludes the following:

> In sum, more and better research will be needed to determine if, as some claim, input is relatively unimportant in L2 learning. To adequately assess the role of L2 input, the input that learners of an L2 actually receive must be assessed more accurately. Measuring L2 input may be impossible, but better estimates of L2 input can and must be obtained. Doing this will require the expenditure of substantial resources (time, money, creativity). For this to happen, researchers must first decide to give L2 input a chance to explain variation in L2 learning.


Studies on miniature artificial languages, often statistical in nature, are designed to control as many variables as possible. Depending on the focus of the study, features of the artificial language may or may not be permissible in the participants’ L1 and may be phonological, morphological, syntactic or semantic. Experiments are conducted in order to test the learners’ performance in the novel language, in particular with respect to these pre-selected features and relative to variables such as working memory, rule learning, generalization, knowledge representation, and other processing and cognitive functions (cf. for example Hudson Kam & Newport 2005, Reber 1967, Saffran et al. 1997).
Although the findings of this type of research are often relevant to research in language acquisition and teaching, there are clear limitations. Johnson (2006) speaks about fundamental differences between real languages and artificial languages and argues that in order to make them learnable in a short span of time, miniature languages necessarily lack the complexity and variation seen in natural languages. The high degree of control comes at a cost of ecological validity.

The alternative is a close examination of the natural linguistic input to which learners are exposed. Few studies have attempted to do this, primarily because of the difficulty involved in controlling, measuring, coding and documenting the language environment of the learner. Recently, however, a few studies have begun to investigate ways to observe and examine the first stages of TL acquisition with a complete control of the linguistic input.

- Rast’s (2003) doctoral work conducted in France on French learners of Polish in the first 8 hours of acquisition showed that salience significantly helped learners perceive and comprehend items in the Polish input.
- Hendriks & Prodeau (1999, 2000) looked at the acquisition of Dutch finite verb placement in French learners of Dutch with or without prior knowledge of German as a foreign language. The authors found that, whereas all learners received the exact same input in which evidence for the verb placement was sparse, only those with prior knowledge of German discovered the target verb placement rule.
- Gullberg, Roberts, & Dimroth (accepted) found that Dutch learners exposed to video typed Chinese input for only 7 minutes were helped in word recognition and sound to meaning mapping by a combination of item frequency and gestural highlighting.
- Winkler (2010) ran an intervention study in which two groups of Italian L2 learners of German were exposed to 60 hours of classroom teaching each. The input that was entirely recorded mainly differed in the orders in which certain grammatical features were introduced. The manipulation had dramatic effects on the learners' performance.

These studies, however, have only touched the surface in terms of our knowledge about the psycholinguistic and cognitive processes involved in language appropriation. Measuring natural language input is not impossible as is evidenced by the studies mentioned above. It does, however, present new methodological challenges, as Flege (2009) rightly points out. This project is designed to meet these new challenges in a set up that is close to real life.

3 The theoretical and methodological basis for the work and why this will be feasible, innovative and interesting

The current project proposes a cross-linguistic empirical study within the functionalist theoretical framework with both theoretical and applied objectives. Variables will be controlled in order to respond to the following questions:

1. What is a learner capable of perceiving, comprehending, and producing in a novel TL at first exposure and within the first 20 hours that follow? What processes are active even before learners are able to express themselves in the TL by means of non-formulaic constructions?

2. What is the role of the SL and other foreign languages previously studied in these processes? What knowledge do learners make use of, independent of their SL and other L2s, such as discourse principles, pragmatics and information structure to acquire a new language? What extra-linguistic (world) knowledge do learners rely on to appropriate a TL?
3. What is the role of the linguistic input itself (frequency of items and features in the input, for example)? Which paralinguistic cues (gestural, interaction, context) are used?

4. How do learners’ language abilities develop over time? Across proficiency levels, do learners shift their attention from semantic to structural aspects or do they gradually develop both of them?

5. What are the effects of different types of presentation of similar target language properties (mere exposure vs. availability of more explicit information on the structure of the language)?

6. How do children and adults compare with respect to extracting information from the input?

In order to address these questions the project proposes a complete documentation of both the instructor’s and the learners’ productions (speech and written texts) in the classroom. The syllabus will consist of a mixture of prefabricated building blocks (e.g. written texts, recordings of dialogues, short video clips etc.) that will be the same for all learner groups and the semi-spontaneous oral input produced by the instructor. The oral input will be scripted and kept constant to the degree possible for online speech production in typical classroom interactions. Care will be taken to make sure that the input contains the lexical elements needed by the learners for coping with the language production tasks included in the test series.

For the "Explicit Input" condition (question 5) additional materials (e.g. tables displaying inflectional paradigms) and other methods of visualizing linguistic relationships and dependencies will be employed. The prefabricated elements and the scripted oral input for the adult groups ("Basic Input" and "Explicit Input") will be designed such that they can also be used with the 10-year old learners in the "Child Learner" group (question 6), even though slight adaptations of the style of presentation might be unavoidable.

If we wish to understand how learners work on their TL input, what in the input they “take in” (attend to, notice, perceive, etc.) and how they take it in, we need evidence of this activity. By means of quality recording and filming of classroom interactions with appropriate technologies, we can control for many more aspects than has been feasible in the past. The instructor’s and the learners’ speech will be recorded separately. In this way, the researchers will have the opportunity to analyze learner performance with respect to the instructor input, the learner input, or both. Professional equipment is necessary to guarantee high-quality recordings so that learners’ analyses of various aspects of the Polish language from phonology and morpho-syntax to pragmatic organisation can be carefully studied (hence the need for a multi-entry recording apparatus with compatible microphones that will be jointly used by the collaborating partners). Perception, repetition and comprehension tasks will be conducted at various time intervals during the duration of the Polish instruction and production data will be elicited as early as possible (hence the need for several Polish native speakers to transcribe, code, and translate data).

Polish was chosen as the target language primarily because of its structural properties and its contrasts with the source languages of the study. Polish has rich morphology, free word order, pro-drop, no obligatory determiner, and a temporal-aspectual system in which aspectual marking is independent of verbal tense marking. Our learners will be at a stage where they are unlikely to produce such markers yet; however, it will be interesting to see how they interpret the meaning of such markers in the input.
Analysis at the phonological level also points to interesting contrasts. Polish, for example, allows more complicated consonant clusters than any of the source languages of the study. As for stress placement, it will be interesting to compare the acquisition of Polish by speakers of languages with lexical stress (English, Dutch, German) vs. fixed stress (French), where the stress falls, however, in a different place than in Polish. Will these contrasts affect our learners’ acquisition of Polish differently relative to their L1, and if so, in what ways?

The choice of Polish has also practical advantages. Novice learners are potentially abundant in the participating countries where Polish is rarely taught. The four countries of the study also have large immigrant populations, so knowledge of Polish and the experience of learning Polish will be useful for future language teachers. Most PIs in the project have at least basic competence in Polish.

4 Other relevant published and continuing research on which it will build

Several strands of prior and ongoing research feed into the research proposed in this project. The development of the project's methodology profits greatly from Rast’s dissertation published in Rast (2008) and subsequent work focusing on specific areas of early L2 acquisition, such as how learners convert TL input to “intake” (Rast 2010), how learners make use of linguistic knowledge other than their L1 to appropriate a new foreign language (Rast, forthcoming-a), and how first exposure data can contribute to models of input processing as in VanPatten’s (2004) principles, the Primacy of Content Words Principle and the Sentence Location Principle (Rast, forthcoming-b). Additional pilot studies aiming at a refinement of the input structure and trying out different presentation modes are currently under way (U. Paris 8: Rast & Watorek).

The set-up for the current study is also informed by results from a second strand of L2 acquisition studies (Gullberg et al., accepted; Zwitzerlood et al. 1994) that undertake the study of learners’ processing of natural language input at the initial stages of L2 acquisition but sacrifice the relative naturalness of the learning situation in favour of more control. In these studies, every learner is exposed to relatively small amounts of the exact same foreign language input (videotaped). There is no interaction, participants are not even asked to try and learn something, and still they can be shown to have acquired some information about word forms and meanings in the new language after as little as 7 minutes of exposure (Gullberg et al., accepted). Through careful manipulations of the input, this approach allows to study the effect of selected input properties under the magnifying glass, a technique that can be applied to study the effects of some of the pre-fabricated input units. When it comes to learners' comprehension of written input, we can also build on recent findings by Peyer, Kaiser, & Berthele (to appear) concerning the comprehensibility of text passages by learners with no or little target language knowledge.

The part of our research concerned with distinct approaches to presenting the same target language input is inspired by recent criticisms of the communicative approach to language teaching. Puren (1995) calls into question methods that focus solely on situations, communication and/or speech acts and for which eclecticism has potentially contributed to a decrease in teacher investment. This concern triggered further theorizing about the nature of the communicative approach and the need to provide learners with more information about the language than what they were receiving, much like studies of focus on form (Doughty & Williams 1998), input modification (Faerch 1986, Long 1983), and input enhancement (Sharwood Smith 1993), as well as the Processing Instruction model proposed by VanPatten, (2004). For recent studies comparing classroom acquisition under different input conditions
The proposed investigation concerning the age factor can add to an impressive amount of research (cf. Singleton & Ryan 2004, for an overview). This research has, however, been dominated by studies addressing the critical period hypothesis (cf. Birdsong 1999, 2006). Because this hypothesis only makes predictions for ultimate attainment, studies on the initial stages of child second language acquisition are still scarce. Even fewer studies have attempted to compare children and adults learning a second language in a controlled learning/teaching situation (cf. Lazarova’s 2004 PhD for an exception). Children are generally assumed to outperform adults in second language learning, but even after many years of intense research it remains unclear if they are just more “open” to new ways of expressing information than adults, more able to distinguish sounds of the L2 that are in contrast to the L1, better at detecting statistical systematicity, etc. and if differences between children and adults are qualitative or quantitative in nature. Researchers are also far from agreement on the explanation for such differences, that have been ascribed to causal factors as diverse as the type and amount of prior knowledge brought to the task (incl. the depths of ‘L1 entrenchment’), general cognitive maturity, the accessibility of innate language specific constraints (UG), and differences in the input (!).

Our study can build on the findings of some adult-child comparisons this have attempted to keep that last point constant: Hudson Kam & Newport, 2005 use relatively natural exposure to an artificial language in order to test how children and adults deal with uninterpretable variation. Roberts, Gullberg, & Dimroth (in prep.) compare children and adults performance after 7 minutes of exposure to input in a natural language (Mandarin Chinese).

Finally, the general perspective adopted for our study draws on the Leaner Varieties approach (Klein & Perdue 1997). This means that the result of the initial phases of acquisition - elementary and simple as it might be - will be considered as a nascent linguistic system in its own right and not as an imperfect replication of the target language. Only in this way, we believe, can its developing internal structure be uncovered and related to the input in a meaningful way.

With the question, what learners can and cannot do with the linguistic input surrounding them, this project addresses a hotly debated topic in the field of L2 acquisition. By relying on controlled but natural input, and closely observing the learners achievements over time the partners in this 4-lateral collaboration will contribute new empirical evidence to this debate about the very first phases of foreign language acquisition and the nature of elementary linguistic organisation while at the same time providing applicable insights with a clear benefit for classroom research and language pedagogy.

5 The specific research proposed and how this will be contributed to by the various national partners, the research timetable, the resources required, and the contributions of the various team members

In order to detail precise information about how this project will be conducted, an extensive table (Table 5.1) provides an overview of the project and details the dates, procedures, teams, and countries where costs will be supported (the table appears in the Appendix).
Individual country reports provide information about personnel to be hired for the project, PhDs, Post-Docs, research engineers, research assistants, student learners, and a Polish instructor, and specifics about the equipment and resources needed. The details of the Polish course and data collection are presented in Section 6 of this document, where it will become evident why these positions are all necessary for the ultimate success of this project.

With regard to the focus of various team members and areas of investigation, this project provides the opportunity to investigate input processing of a novel language at various levels (phonology, lexicon, morpho-syntax, pragmatic utterance organisation) as well as to conduct analyses across themes (L3 acquisition, individual differences, and the acquisition-teaching interface). Collaboration between researchers participating in the project will consist of sharing and comparing their results in different areas. In fact, the strength of this project lies in this opportunity to exchange results across diverse areas of investigation with a view to constructing theoretical objects that originate from shared questions (the L2 initial state, the age factor, the acquisition-teaching interface) as well as applied objects, such as implications of empirical studies on language acquisition and concrete pedagogical procedures. Table 5.2 (see the Appendix) indicates the specific areas of investigation of project researchers.

Members of the project will collaborate as well with external partners in Italy, Switzerland and the Netherlands (see details in Appendix Section 5).

6 Details of any specifics about evidence collection and of the proposed evidence analysis, and why these are relevant, appropriate and innovative

Decisions about the design of the Polish course and the methods of data collection are primarily based on an assessment of methodologies used in three prior experiments on French learners of Polish conducted at the University of Paris 8 (Rast 1998, 2003 and a current pilot study). Three types of data will be collected: 1) instructor input; 2) learner input; 3) tests. The two types of input (instructor and learner) will be recorded separately so that during the data analysis phase, the two types of input may be combined or analyzed separately. The instructor and students will be recorded by means of professional equipment to allow for analysis of their productions not only at the level of the utterance with regard to morpho-syntax and lexical usage, but also at the phonetic and phonological levels. These data will be transcribed, translated into English and into the L1 of the particular learner group for contrastive analysis purposes, and coded.

All learner groups of the project will take the same tests. Polish native speaker control groups will also take these tests (data collection in Poland). In some cases, when we wish to identify what a learner may be able to do before exposure to the Polish language, we will administer a given test to a group of native speakers (French, Dutch, German, English) who have no knowledge of Polish and who are not learners of Polish, referred to as “first exposure” participants. The development, piloting and modification of tests is currently underway and will continue through Fall 2010. By Spring 2011, we will have an established battery of tests to be administered to the participants of the current study.

The Polish course will take place over 10 class sessions of 120 minutes each with 15 learners per group. In France and the Netherlands, two input exposure types will be presented to two groups of the same type of learner (adults - typically university students). In the UK and Germany, the same type of input will be presented to a group of children (10 year olds) and a group of adults (see Table 6.1).
Table 6.1. General overview of the data to be collected

<table>
<thead>
<tr>
<th></th>
<th>France</th>
<th>Netherlands</th>
<th>UK</th>
<th>Germany</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic input</strong></td>
<td>20h/15 subjects</td>
<td>20h/15 subjects</td>
<td>20h/15 subjects</td>
<td>20h/15 subjects</td>
<td>80h/60 subjects</td>
</tr>
<tr>
<td><strong>Explicit input</strong></td>
<td>20h/15 subjects</td>
<td>20h/15 subjects</td>
<td></td>
<td>40h/30 subjects</td>
<td></td>
</tr>
<tr>
<td><strong>Child learners</strong></td>
<td></td>
<td>20h/15 subjects</td>
<td>20h/15 subjects</td>
<td>40h/30 subjects</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>40h/30 subjects</td>
<td>40h/30 subjects</td>
<td>40h/30 subjects</td>
<td>40h/30 subjects</td>
<td>160h/120 subjects</td>
</tr>
</tbody>
</table>

Materials for the Polish course are currently being developed, piloted and modified. The same materials will be used for all learner groups, with the exception of the “Explicit input” group, which will receive additional information about formal rules of the TL. Themes will consist of typical topics used in communicative-style beginning language instruction: introductions, family, food, restaurant, travel, directions, etc. (see the Protocol for the Polish course in Appendix).

With regard to the instructor’s input, the following variables will be controlled:

- **Lexicon**: Lexical items will be selected and taught relative to certain tasks the learners will be asked to perform.
  - The transparency of a given lexical item will be controlled for all L1s of the study. Transparent items are those Polish items that share phonetic and orthographic features with an item in the L1. A database of lexical items that are transparent between the languages of the study is currently being developed.
  - Frequency (see e.g. Rott 1999) of lexical items in the input will be controlled and measured by means of the CHILDES tools (MacWhinney 2000). A database of “frequent” items is being established based on prior Polish input studies and will be adapted according to the final Polish instruction procedures.

- **Structure**: Structures will be controlled for frequency in the input.
  - Nominal morphology will be controlled and analyzed with respect to the system of nominal morphology in Polish and that in the L1s of the study.
  - Word order will be controlled and analyzed with respect to word order in Polish and word order in the L1s of the study.

Two variables will be held constant for the two types of input and all tests: transparency and frequency of an item in the input. Statistics will be collected to reveal inter- and intra-group variance and main effects of the variables across groups of learners and over time.

Testing will follow a longitudinal design; each test will be administered three times during the experimental period. The tests that focus learners’ attention will be administered at the end of a class session. Tests will be adapted appropriately for children. Sample areas of investigation are perception and segmentation, perception and comprehension, comprehension of verbal utterances relative to word order and morphology, and written and verbal production. Non-linguistic psychometric tests will also be administered to control for individual differences (cf. Appendix Section 6 for more details about these tests).
7 Ethics: Any ethical issues arising from the research, how they will be addressed and that they have been subject to all appropriate national and institutional procedures.

Each team will respect the ethical rules of its respective country, in terms of authorization by the relevant ethical committees, as well the confidentiality and anonymity rules and the restrictions relative to computer-generated personal files. The project involves data collection from human participants who are unable to give informed consent (children). Each team that works with such populations will apply for authorisation from the relevant authorities.

8 Data management: Arrangements made for data storage and providing access to data for other researchers

Data will be archived in the Browsable Corpus Archive hosted at the MPI for Psycholinguistics in Nijmegen (http://www.mpi.nl/resources/data). Prof. Wolfgang Klein (Director) has agreed to this solution.

9 Staff duties: What are the roles of the various staff proposed on the project?

The chart below portrays the internal synergy of the project and indicates the composition of the Steering Committee and individual responsibility for the different tasks and operations of the project. Given the size of the French team, it was agreed that two researchers from France should serve on the Steering Committee. Table 5.2 (see Appendix) shows the distribution of tasks with respect to data analysis and the valorisation of results.
France  Watorek/Rast
Netherlands  Starren
Germany  Dimroth

Steering Committee

1st period of data collection  FR/NL/UK/GE (Basic Input)
2nd period of data collection  FR/NL (Explicit Input), UK/GE (Child Learners)

Data acquisition

Empirical investigations
- Phonology  Shoemaker
- Lexicon  Dijkstra
- Morpho-syntax  Lenart
- Pragmatic organization  Hendriks
- Individual differences  Hilton

Analysis
- Initial state in SLA  Rast
- Interface teaching/acquisition  Véronique
- L3 acquisition  Trévisiol
- Factor age  Dimroth

Theoretical topics

Practical implications for foreign language teaching  Starren

Valorisation
- 2 General meetings (NL and GE)
- International conferences (all teams)
- Publications (all teams)
- Project final conference (France)
10 Impact: Who are seen as the beneficiaries of the research, both academic and non academic impact and how this research will be disseminated to them?

This project will be, without a doubt, of interest to the academic community in that it is based on an extensive empirical study of rigorously prepared and controlled linguistic input that will contribute significantly to current debates and discussions focused on important theoretical questions, such as the L2 initial state, the acquisition-teaching interface, and L3 acquisition. The methodology proposed in the study is innovative and allows for the control of a multitude of factors. Researchers working on language acquisition, input factors, or classroom teaching and interaction will benefit methodologically from this project. In addition, a better understanding of child and adult L2 acquisition in relation to controlled input will benefit language acquisitionists in general. The project is a leap forward from acquisition studies without control over the input (i.e., most current studies) and from studies that investigate only child L2 or adult L2 acquisition (i.e., that do not carefully control for age). The results should shed light on the age factor more generally. Given that the project is cross-linguistically oriented, it also carefully controls for not only universal cognitive factors such as age, but also for language-individual differences. Researchers interested in the debate regarding the language/cognitive development interaction and its influence on language acquisition will therefore also benefit from our results.

The project will also be of interest to the academic community on a more practical level as it proposes to work within the context of teacher training programs, using as its base the structure of a course developed in France entitled Learning and Self-Observation (see Appendix Section 10 for details). In this way language teachers and students at the primary, secondary and tertiary levels will also benefit.

A principal objective of this project, however, is to address the urgent issue of helping render language teaching methods and language learning as efficient as possible so that today’s students become competent users of their L2s. The question of how to render language learning more efficient is not new, but rather, recurrent and has been at the core of studies for decades. The manner in which these studies have tried to respond to this question, however, has often been too general, resulting in less precise information than what is needed. The intellectual and compelling reflections on relations between language acquisition and teaching based on large quantities of acquisition studies often fail to provide concrete assistance to language teachers or learners. This project provides the perfect opportunity to bring together the worlds of language pedagogy and acquisition in that the data collected will describe what a learner progressively does to appropriate an L2 in an instructional setting, as articulated by Watorek (2008). We will then be able to respond to very precise questions, such as what helps a learner perceive elements of the input, what type of presentation of morpho-syntactic information better increases a learner’s chance to appropriate the information, and how this differs for children and adults.

Dissemination of research

The data will be available in a browsable database by the end of the project (December 2013). The data should be fully transcribed by end of 2012. Data can then be converted into the format needed for the Max-Planck browsable database, and other databases.

All publications and conferences regarding the data are planned for the year 2013 in which analyses should be largely complete. Presentations are planned at conferences with academic topics in the areas of second language acquisition such as the EUROSLA 2013 conference, the Cognitive Linguistics conference, the next AILA conference. The results will also be
presented at conferences with a narrower focus on foreign language teaching, such as the ICC International Conference and the ALTE international conference.

Publications again will be submitted to journals appropriate for the academic side, i.e., research in second language acquisition (SSLA, Language Learning, LIA), and journals appropriate for the teaching and language policy related side such as IRAL. The Taylor & Francis Group of Routledge has contacted one of the authors with an invitation to submit a manuscript to Alison Mackey, co-editor of the Second Language Acquisition Research book series.

We propose to host a website regarding the project, which can be visited by all interested parties and which contains the links to the browsable database and other databases. The website will also include information about the project’s development.

All subjects will receive a certificate of attendance, and will be briefed about the project once they have participated.

The final project conference will actively seek to invite academics, language policy makers, teachers, and possibly even learners of second languages.

References


Hendriks, H., & Prodeau, M. (2000). ‘Acquiring new syntactic patterns: “But the teacher never told me that French and Dutch word-order were different!”’ EUROSLA 10, Kraków, Poland, September.


Roberts, L., Gullberg, M., Dimroth, C. (in prep.). What can adults and children learn about word forms in ten first few minutes of exposure to a new language?


Winkler, S. (2010). 'Learner varieties and language teaching curricula. An intervention study on the acquisition of German (S)VO clause structure by native Italian speakers'. Project meeting *The Structure of Learner Varieties*, Max-Planck-Institute for Psycholinguistics, Nijmegen, March.


Appendix
for

Varieties of Initial Learners in Language Acquisition (VILLA):
Controlled classroom input and elementary forms of linguistic organisation

Section 5

Table 5.1 The project organization

<table>
<thead>
<tr>
<th>Dates</th>
<th>Procedure</th>
<th>Teams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.-Mar. 2011</td>
<td>Recruiting of personnel, preparation of Polish course</td>
<td>French teams + Polish teacher</td>
</tr>
<tr>
<td></td>
<td>First meeting of all participants in France (coordination meeting 1)</td>
<td>All teams</td>
</tr>
<tr>
<td>Apr.-June 2011</td>
<td>Pilot project, training of Polish instructor, test preparation and piloting of tests</td>
<td>French teams + Polish teacher, research engineer, 2 hires on a sessional basis for transcription, glossing, and coding of data</td>
</tr>
<tr>
<td></td>
<td>Piloting of tests</td>
<td>All teams</td>
</tr>
<tr>
<td></td>
<td>Coordination meeting (2) in Germany</td>
<td>All PI</td>
</tr>
<tr>
<td>July-Sept. 2011</td>
<td>Constitution of participant groups in the 4 countries</td>
<td>All teams</td>
</tr>
<tr>
<td>Oct.-Dec. 2011</td>
<td>First experimental phase: data collection and test results (4 base groups). Transcription</td>
<td>All teams Polish teacher based in France spend 3 weeks in each country</td>
</tr>
<tr>
<td></td>
<td>Coordination meeting (3) in France</td>
<td>All PI and French Post-Doc</td>
</tr>
<tr>
<td></td>
<td>Data collection from control group of Polish natives speaker in Poland</td>
<td>French researchers</td>
</tr>
<tr>
<td>Feb. 2012</td>
<td>Transcription (continued) first analyses</td>
<td>All teams Creation of a database coordinated by the research engineer (France) in collaboration with MPI</td>
</tr>
<tr>
<td>Mar.-June 2012</td>
<td>First General meeting – all participants, presentation of first analyses (valorisation)</td>
<td>Meeting in NL</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Second experimental phase: data collection:</td>
<td>NL: contract for the Polish instructor</td>
<td></td>
</tr>
<tr>
<td>- for the variation of type of input corpus (NL/France)</td>
<td>France/NL Polish instructor’s stay for 3 weeks in NL and in France</td>
<td></td>
</tr>
<tr>
<td>- and variation of the type of learner (UK/Germany).</td>
<td>Germany/UK Polish instructor’s stay for 3 weeks in Germany and NL</td>
<td></td>
</tr>
<tr>
<td>Beginning of data transcription</td>
<td>France: 2 hires on a sessional basis</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>July-Sept. 2012</th>
<th>Continuation of transcription and first analyses</th>
<th>All teams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordination meeting (4) in UK</td>
<td>All PI</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oct.-Dec. 2012</th>
<th>All data will be transcribed and glossed</th>
<th>All teams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complementary collection of data from control group of Polish speakers</td>
<td>French researcher</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jan.-Mar. 2013</th>
<th>Work on data – meetings, writing, dissemination (participation at conferences)</th>
<th>All teams</th>
</tr>
</thead>
<tbody>
<tr>
<td>General meeting (2) all groups in Germany (preparation of the final conference)</td>
<td>Participation of all researchers of the project</td>
<td></td>
</tr>
</tbody>
</table>

| Apr.-June 2013 | Work on data – meetings, writing, dissemination (participation at conferences) | All teams |

<table>
<thead>
<tr>
<th>July-Sept. 2013</th>
<th>Work on data – meetings, writing, dissemination (participation at conferences)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordination meeting (5) in NL</td>
<td>All PI</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oct/dec 2013</th>
<th>Final conference in Paris</th>
<th>All teams + invited speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication of collective volume</td>
<td>All teams</td>
<td></td>
</tr>
</tbody>
</table>

The information presented in this table corresponds to the figures presented in the “Cost Summary” and “Justification of resources”

Table 5.2 Areas of investigation

<table>
<thead>
<tr>
<th>Source language</th>
<th>Phonology</th>
<th>Lexicon</th>
<th>Morpho-syntax</th>
<th>Pragmatic utterance organisation</th>
<th>L3 acquisition</th>
<th>Individual differences</th>
<th>Interface teaching/acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoemaker post-doc</td>
<td>Rast Hilton Lenart Véronique Watorek Trévisiol Rast</td>
<td>Hilton post-doc</td>
<td>Véronique Watorek</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source language</td>
<td>Shoemaker</td>
<td>Rast</td>
<td>Hilton</td>
<td>Hendriks</td>
<td>Hendriks</td>
<td>Hendriks</td>
<td>Hendriks</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>English</td>
<td>Lenart</td>
<td>Paprocka</td>
<td>Lenart</td>
<td>Watorek</td>
<td>Post-doc</td>
<td>Post-doc</td>
<td>Post-doc</td>
</tr>
<tr>
<td>Dutch</td>
<td>Hendriks</td>
<td>Starren</td>
<td>Hendriks</td>
<td>Starren</td>
<td>Starren</td>
<td>Starren</td>
<td>Starren</td>
</tr>
<tr>
<td>German</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
</tr>
<tr>
<td>TL Polish</td>
<td>Post-doc</td>
<td>Lenart</td>
<td>Post-doc</td>
<td>Paprocka</td>
<td>Watorek</td>
<td>Post-doc</td>
<td>Post-doc</td>
</tr>
<tr>
<td>INPUT variations (FR and NL)</td>
<td>Shoemaker</td>
<td>Rast</td>
<td>Hilton</td>
<td>Lenart</td>
<td>Watorek</td>
<td>Véronique</td>
<td>Trévisiol</td>
</tr>
<tr>
<td></td>
<td>Hendriks</td>
<td>Lenart</td>
<td>Hendriks</td>
<td>Paprocka</td>
<td>Watorek</td>
<td>Hendriks</td>
<td>Hendriks</td>
</tr>
<tr>
<td>Age (UK and GE)</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Dimroth</td>
<td>Hendriks</td>
<td>Hendriks</td>
<td>Hendriks</td>
<td>Hendriks</td>
</tr>
</tbody>
</table>

**External collaborators**

Members of the project will closely collaborate with the following external partners:
Giuliano Bernini (Università di Bergamo, IT)
Raphael Berthele (Université de Fribourg, CH)
Leah Roberts (MPI for Psycholinguistics, Nijmegen, NL)
Steffi Winkler (Vrije Universiteit Amsterdam, NL)

Giuliano Bernini and Raphael Berthele proposed to contribute to a potential extension of the project by applying for additional grants from national funding agencies in Italy and Switzerland. This would allow to reuse the methods (input structure, test batteries) developed in this project in order to obtain comparable data for learners with different L1s and for learners with prior knowledge of one or more source languages (+/-multilingual).

**Section 6**


<table>
<thead>
<tr>
<th>Session</th>
<th>Input</th>
<th>Cumulative Input</th>
<th>Themes</th>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2h00</td>
<td>2h00</td>
<td>« Presentations »</td>
<td>- lexical decision</td>
</tr>
<tr>
<td>Time</td>
<td>2h00</td>
<td>4h00</td>
<td>6h00</td>
<td>8h00</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>- members of family, ages, food, drink, nationalities, professions, languages, countries</td>
<td>- eat, drink, fruits, vegetables, desserts, etc.</td>
<td>- vocabulary review</td>
<td>- Where is... ? prepositions, places – train station, airport, left, right</td>
</tr>
<tr>
<td>3</td>
<td>« word order (written) »</td>
<td>« segmentation »</td>
<td>« lexical decision »</td>
<td>« word order (written) »</td>
</tr>
<tr>
<td></td>
<td>« judgements (written) »</td>
<td>« psychometric tests »</td>
<td>« grammaticality judgement-morphology »</td>
<td>« judgements (written) »</td>
</tr>
<tr>
<td>4</td>
<td>« translation of written tests »</td>
<td>« comprehension/syntax »</td>
<td>« translation of written tests »</td>
<td>« translations of written tests »</td>
</tr>
<tr>
<td>5</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
</tr>
<tr>
<td>6</td>
<td>« translation of written tests »</td>
<td>« comprehension/syntax »</td>
<td>« translation of written tests »</td>
<td>« comprehension/syntax »</td>
</tr>
<tr>
<td>7</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
</tr>
<tr>
<td>8</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
</tr>
<tr>
<td>9</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
</tr>
<tr>
<td>10</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
<td>« comprehension/syntax »</td>
<td>« psychometric tests »</td>
</tr>
</tbody>
</table>

The following is a list of areas of investigation and a *selection* of tests proposed to collect data to respond to our questions posed in Section 2. The names of the three principal researchers involved are also given. Other tests are currently being discussed and developed.

1) Perception and segmentation
a) **Lexical decision test** (Rast, Shoemaker, Watorek)

*Research questions:* What phonotactic knowledge do learners come to the acquisition process with that could be useful when confronted with a new language? What helps learners recognize certain features in the input, in this case the phonotactics of the Polish language?

*Hypothesis:* First exposure participants will perform at chance level. After a few hours of instruction, learners will be able to recognize Polish words, in particular those that are transparent and/or frequent in the input.

*Participants:* First exposure group, learners, Polish native speaker controls

*Description of test:* Randomized Polish and Russian words are recorded by a bilingual Polish-Russian speaker. The Russian words selected do not adhere to the phonotactic rules of Polish. Participants are asked to listen to a word and decide if it is Polish or not.

*Analysis of results:* Using E-prime, this test will measure accuracy and reaction time with respect to the variables of frequency in the input and transparency.
b) **Segmentation test** (Piotrowski, Rast, Shoemaker)

*Research questions:* How does a learner break into the speech stream of an unknown language? What helps the learner segment aspects of the speech stream?

*Hypothesis:* Transparency, frequency in the input and position in the utterance are variables that will affect how a learner perceives and segments items in the input.

*Participants:* First exposure group, learners, Polish native speaker controls

*Description of test:* Sentences are randomized. Each sentence appears three times in the test, once with the target word in initial position, once in medial position, and once in final position. Target words are selected with respect to transparency and frequency in the input. Participants listen to a sentence followed by a word and decide if the word was in the sentence they heard. Distractors are Polish words that do not appear in the test.

*Analysis of results:* Using E-prime, this test will measure accuracy and reaction time with respect to the variables of frequency in the input and/or transparency.

---

2) Perception and comprehension of words

**Lexical transparency test** (Dimroth, Lenart, Starren)

*Research question:* At first exposure, what words in the TL will learners be capable of comprehending and translating simply because they resemble a word in their L1?

*Discussion:* Generally we imagine these words to be cognates; however, prior research has shown that words considered cognates from an etymological perspective are not always perceptible to learners as such. For example, in Rast’s study (2003), the Polish word *koleżanka* (colleague; collègue in French) was only translated correctly by two (or 3%) of the 34 French participants, whereas Polish *uniwersytecio* (“université” in French) was translated by 74% of the participants. This indicates that the linguistic analysis of cognates is insufficient when examining the psycholinguistic processes of perception and comprehension. We need to know what learners themselves perceive as close or distant between their L1 and the TL (cf. psychotypology, Kellerman 1983). For this reason, we designed the lexical transparency test to identify which words in the Polish input are inherently “transparent” for our learners.

*Hypothesis:* The higher the phonetic and orthographic similarity of the Polish word is to a word in the learner’s L1 that carries the same meaning as the Polish word, the higher the success rate of translation will be.

*Participants:* First exposure participants, learners.

*Description of test:* In order to develop a database of words that are transparent across languages of the study, we include a core of Latin-based words in the test. Words used in other tests will be included as well in order to control for the variable “transparency” across tests and language activities (perception, comprehension, grammatical analysis, production). Two versions of the test are administered to two different groups, a written and aural test. A randomized list of Polish words in isolation is presented in list form for the written condition and recorded for the aural condition. Participants are asked to read or listen to the words and to translate them as best they can into their L1.

*Analysis of results:* According to percentages of correct translations, words will be categorized as “opaque” (rating of 0% correct translations), “fairly transparent” (rating of 1-50%) and “very transparent” (rating of 51-100%) to be used for further analyses of the Polish input and learners’ performance.

---

3) Comprehension of verbal utterances

a) **Grammaticality judgement test for morpho-syntax sensitivity** (Lenart, Rast, Watorek)
Research question: The ESF project (Perdue 1993) found that morphological markers are rarely produced in initial learner varieties, but are they perceived? If so, how do learners manage to generalize morphological rules in Polish? Children have generally been found to be more successful with morphology than adults. Will we find this in the early stages of acquisition as well?

Hypothesis: We hypothesize that with minimal input, learners of all groups will perform at chance level. It is only with a certain quantity of input (unknown at this point) that learners will begin to be sensitive to nominal morphology and be able to generalize the rules, and that children learners will develop in this area more quickly than adults.

Participants: Learners, Polish native speaker controls.

Description of test: Sentences are designed to control for case and gender. Some sentences show correct forms for case and gender, and others incorrect. Participants are asked to listen to randomized sentences and to decide if the sentence is correct or incorrect.

Analysis of results: Using E-prime, results will be analyzed with respect to accuracy and reaction time.

b) Comprehension morpho-syntax test (Dimroth, Rast, Watorek)

Research questions: To what extent does syntax play a role in the comprehension of utterances? How do learners analyze inflectional morphology with respect to word order in a highly inflected language with flexible word order like Polish?

Discussion: The Polish system of flexible word order allows for all possible word order combinations. The following sentence is possible in Polish:

\[
\begin{array}{ccc}
\text{matkę} & \text{kocha} & \text{syn} \\
\text{Acc} & \text{Vf} & \text{Nom} \\
\text{mother} & \text{loves} & \text{son} \\
\end{array}
\]

\(= \text{The son loves his mother}\)

For learners who are accustomed to relying on word order to find information about agent and patient, this type of sentence could pose a problem for comprehension for learners whose L1 is SVO.

Hypothesis: We hypothesize that learners who have not yet acquired nominal morphology will adopt the word order strategy of their L1 to process a sentence. Such learners are predicted to incorrectly interpret the sentence in the example above as “The mother loves her son”. Native speakers of German may have an advantage over native speakers of French, Dutch and English because of a greater degree of case marking in German.

Participants: Learners, Polish native speaker controls.

Description of test: Sentences with transitive verbs and two animate agents (e.g. the dog pulls the woman) are randomized and recorded. Each sentence is matched with a picture that illustrates the scene. All scenes have the inverse sentence and picture (e.g. the woman pulls the dog). Participants listen to a sentence and see two pictures. They are asked to match the correct picture with the sentence.

Analysis of results: The eye-tracking apparatus will allow for analysis of accuracy, reaction time and eye movement during the experiment.

4) Comprehension and Production

a) Written word order test (Komur, Trévisiol, Watorek)

Research question: Do learners rely on the syntax of their L1, on that in the input, or on other factors (e.g. the syntax of other known L2s, pragmatic factors) when producing sentences in Polish?

Hypothesis: Several hypotheses currently exist regarding how learners process syntactic information: 1) Learners will rely on their L1; 2) learners will rely on their L1 or another known L2, whichever is most similar to the TL being learned; 3) learners will rely on
pragmatic organisation, such as “controller first” (agent before patient) or “scope” principles (a word precedes that which it has scope over, e.g. in “not hungry”, “not” has scope over “hungry” and therefore precedes it).

Participants: Learners, Polish native speaker controls.

Description of test: Participants read one context Polish sentence followed by a sentence with scrambled words. They are asked to put the words in the appropriate order in Polish.

Analysis of results: Responses will be analyzed with respect to the word orders produced relative to learners’ L1, other known L2s, and the frequency of particular orders in the Polish input. In particular, we will examine the position of scope particles (e.g. nie – not; również – aussi).

b) Sentence repetition (cf. Schimke 2009) (Dimroth, Piotrowski, Rast)

Research question: Are learners more sensitive to some TL grammar violations than to others?

Hypothesis: L2 learners might be more sensitive to word order violations (in particular in domains with a clear syntax-semantic relation, e.g. the position of negation) than to morphological violations (e.g. case marking).

Participants: Learners, Polish native speaker controls

Description of test: Participants hear grammatical and ungrammatical Polish sentences over headphones and are asked to orally repeat them. Responses are recorded. The sentences are too long to memorize them as chunks, so learners have to reconstruct (and potentially correct) them on the basis of their learner grammar. Processing load is however lower than in completely free production because the repetition condition facilitates lexical retrieval. This allows to test constructions that are slightly more complex than what learners are able to freely produce and it allows to elicit data directly speaking to the phenomena in question (e.g. manipulate the position of verb and negation, or features of nominal morphology).

Analysis of results: Responses will be analyzed with respect to the amount of corrected sentences in each of the domains manipulated.

5) Production (The majority of researchers will be involved in the analysis of these tests, but in particular: Dimroth, Hendriks, Komur, Lenart, Paprocka, Starren, Trévisiol, Véronique, Watorek)

To collect production data, we will exploit well-established retelling tasks such as Charlie Chaplin films, Reksio, The Horse Story, The Cat Story, The Frog Story, or the Finite Story, frequently used in language acquisition studies to collect verbal production data from learners and native speakers. Specific areas of investigation will be pragmatic organisation (the controller vs. the controlled; scope particles) and the development of the noun phrase given that Polish is the only language of the study with no obligatory determiners and highly inflected. We expect the children to have a harder time with the pragmatics of narration than adults; adults should show primarily linguistic problems.

7) Two types of input/exposure

Since we know little about the role of input in learners’ acquisition processes, we know even less about the effect of different types of input on learners in the very first stages of acquisition. Given results of selected studies of input (cf. Piske and Young-Scholten, 2009), we may expect that Basic Input without meta-information will be more effective than the Explicit Input method. In later stages, when expectations for more accurate performance increase, the Explicit Input may be more effective. We will compare results of the different groups on all tests and in all language areas to see if we observe differences, and if so what types of differences. We will also analyze learners’ productions in class with respect to the
type of exposure and observe, to what extent if any, their output differs, and if so, how this might affect their performance on the Polish tests.

8) Learner variability – psychometric-type tests (Hilton, Rast, Trévisiol)
In order to investigate possible roles of individual variables in the initial phases of language acquisition, we will carry out a series of psychometric tests with all learners. Each test to be used has been chosen for its capacity to measure one particular variable relevant to the language-learning process. Ease of administration (and particularly the time involved) has been taken into consideration in the selection of test instruments; recognized psychometric instruments will be used to ensure the validity of the data collected. The variables we will measure range from precise language-learning aptitudes (word memorization, grammatical inferencing, phonological sensitivity), to more general personality traits (extraversion, openness, acceptance of change), and cognitive capacities (memory span, executive function). In all, 13 measures of individual variance will be collected for each learner, with a view to providing complementary information that will help explain some of the variance observed in our acquisition data.

VILLA Psychometric tests for measuring learner variability

<table>
<thead>
<tr>
<th>Variable to be Measured</th>
<th>Test</th>
<th>comments</th>
<th>format</th>
<th>editor</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personality</td>
<td>Alter Ego</td>
<td>Extraversion, anxiousness</td>
<td>coll</td>
<td>ECPA</td>
<td>610.70</td>
</tr>
<tr>
<td>Fluent intelligence</td>
<td>Raven</td>
<td>Light version</td>
<td>coll</td>
<td>ECPA</td>
<td>218.51</td>
</tr>
<tr>
<td></td>
<td>Vocabulary test Binois &amp; Pichot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td>D2 Brickenkamp</td>
<td>Attention test</td>
<td>coll</td>
<td>Hogrefé</td>
<td>105.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>160.80</td>
</tr>
<tr>
<td>Categorisation</td>
<td>Wisconsin Card Sorting Task</td>
<td>Categoric abstraction and strategy</td>
<td>individ</td>
<td>Hogrefé</td>
<td>380.81</td>
</tr>
<tr>
<td>Collocational Lexicon</td>
<td>Verbal automatisms Pluchon Simonet</td>
<td>Memory of pseudowords</td>
<td>coll</td>
<td>Ortho edition</td>
<td>60.00</td>
</tr>
<tr>
<td>Phonological memory</td>
<td>Belec: word perception and phonological memory</td>
<td></td>
<td>individ</td>
<td>ULB</td>
<td>50.00</td>
</tr>
<tr>
<td>Memory</td>
<td>WM Span Daneman &amp; Carpenter</td>
<td></td>
<td>individ</td>
<td>No commercial version exists</td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td>Stroop test</td>
<td>Classic version</td>
<td>individ</td>
<td>To download</td>
<td></td>
</tr>
<tr>
<td>Cognitive preferences</td>
<td>ISALEM-97, Université de Liège</td>
<td></td>
<td>coll</td>
<td>Download from University of Luik</td>
<td></td>
</tr>
<tr>
<td>Morphosyntactic</td>
<td>_Ilama from Meara</td>
<td>Test F</td>
<td>individ</td>
<td>To download</td>
<td></td>
</tr>
<tr>
<td>acquisition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical memory</td>
<td>_Ilama from Meara</td>
<td>Test B</td>
<td>individ</td>
<td>To download</td>
<td></td>
</tr>
<tr>
<td>Phonological sensitivity</td>
<td>_Ilama from Meara</td>
<td>Test D</td>
<td>individ</td>
<td>To download</td>
<td></td>
</tr>
</tbody>
</table>

TOTAL PRICE: 1,586.09 Euro
Section 10

Description of the course *Apprentissage et auto-observation* (*Learning and Self-Observation*)

The experience of learning a new language became an integral part of the training of future language teachers in France in the 1980s. *The Learning and Self-Observation* course, originally introduced within the new created French as a Foreign Language Masters program, allows teacher trainees to adopt the perspective of language students while they explore their own language acquisition experience and observe how different elements and factors come into play in the teaching and learning of a foreign language (cf. Grandcolas & Vasseur 1997). The students, future teachers, choose a language they have never studied before and one that is typologically distant from their native language. They take the language course in parallel with a theoretical course taught by a specialist in language pedagogy. During the theoretical course, students are invited to reflect on their learning experience with the help of texts that theorize pedagogical methods and approaches and to share these reflections in class discussions. These students keep a regular journal in which their writing is guided by self-observation grids on a variety of points, such as the learning situation, pedagogical activities, interactions, learning strategies, and teaching. Our project will serve as a base for this type of experience. For the French side, the Polish courses will be integrated into the curriculum for those training to become language teachers. This same procedure could be developed as well in the institutions of partner countries. Such an experience proves beneficial for training, research and pedagogy and feeds into a reflection about the role of the teacher in enhancing learners’ processes in the construction of new language competence.