GÖTEBORGS UNIVERSITET

Processability and language proficiency: Is there a causal relationship?

Antonio Domenico Saracino

Abstract

The objective of this paper is to investigate if the relationship between the Processability Theory, a model for second language acquisition, and proficiency, is causal. The focus will be on attributive and predicative adjectival agreement. According to the Processability Theory, Swedish is acquired following five steps, which cannot be jumped over. The steps analyzed in this paper are level 3 and 4 of the Processability Theory (PT). In order to investigate this, an experimental study was designed, which consisted of two versions of four texts, identical to each other except for errors in predicative and attributive positions, which put them to different PT levels. The texts were then evaluated by 62 teachers via online survey and the results between the two versions of the texts were then compared. The results clearly show no causal relationship between PT and proficiency. There is no significant difference in mean grade given to the two versions of the texts and, therefore, grade does not depend on PT level.

Keywords: *Proficiency*, *PT*, *processability theory*, *second language acquisition*, *adjective, predicative, attributive*

Table of contents

1. Introduction

In the Swedish school system, teachers make great use of grading matrixes when assessing language proficiencies. These matrixes are though vaguely formulated, and it can be difficult, for a teacher, to grade and assess dozens of texts in the short time required by the schedule. Because of this, the matrixes can be hard to decipher: What is, for instance, the difference between a learner who can express her or himself with *common words* and *simple phrases*, compared to someone who can use *words* and *common phrases*?

The hypothesis behind this paper is bound to the second language acquisition theory known as Processability Theory (PT) and the fundamental question that is here researched is whether there is a direct or indirect relationship between PT and proficiency. This question will, in turn, have bearing on assessment and on the development of the grading matrixes, which will also be discussed.

There is much debate on what language proficiency is and what should be assessed. Some research has described language acquisition in terms of stages: In order to acquire a grammatical structure, a "step" is required to be reached (or "processed") by the learner. One of these theories is the Processability Theory, developed by Manfred Pienemann. Studies have shown positive correlations between processability and language proficiency. However, the question addressed in this work is whether such correlations are causal, that is, if the processability level is connected to a higher grade. The hypotheses are two: 1) null hypothesis, meaning that there is no correlation; 2) that there is, in fact, a relationship. The aim of this study is to verify which one is true.

In order to verify this, an experiment was conducted. Four texts were created and then manipulated into two versions, in total 8 texts, located on different PT levels. These levels are exemplified by attributive and predicative agreement of the adjective which, according to PT, belong to different grammatical competence level. The number of errors is the same in every text which, in turn, also contains several

filler errors. The texts were then graded by assessors, teachers of Swedish as a second language, via online polling, in a scale from 1 to 6. Definitions of the levels in this scale were taken from the Common European Reference Framework for Languages (CEFR).

The predictions are that the grades given to the two different versions of the texts will not differ significantly. The result will, however, have consequences for how we think about the relationship between PT and assessment of language proficiency, and it will be useful to analyze it in the context of specifically assessing grammatical complexity and accuracy.

2. Theoretical background

This chapter is going to describe the theory known as Processability Theory (PT in short from now on). The theory will be presented, followed by a description of the developmental stages for Swedish. I will then describe the Common European Framework Reference for Languages, known as CEFR, used in assessing the learner texts, which will provide the definitions for grammatical accuracy in the scale used to assess the texts created for the present experimental study. Finally, an account of some of the previous research which investigated the relationship between PT and proficiency will conclude the section.

2.1 The Processability Theory

PT originated with the ZISA project, which examined the oral production of 45 learners of German as a second language with Italian and Spanish as L1. PT was introduced after the study to explain the idea of developmental stages in theoretically and coherent way.

PT studies the interlanguage of a learner, which is built on input from the target language and previously learned language, plus internal cognitive processes. It shows only part of the features and structures of the target language, which in turn depend on when, in the learner's voyage to second language acquisition, the "snapshot" of his or her current knowledge of target language (TL) is taken. An interlanguage should therefore be seen as a very dynamic system (Abrahamsson 2009:43). Selinker (1972) coined the term and described interlanguage not as an incomplete language but as a system of its own, that can be investigated and described just like any other language.

PT rises as a theory of second language acquisition which aims to explain developmental sequences in terms of language processing. It is based on a "universal hierarchy of processing procedures that is derived from the general architecture of the language processor" (Pienemann & Håkansson 1999:386). In other words,

language learners can only process structures that are learned in a determined order. These procedures are universal and supported by several studies on the acquisition of different languages, even languages that are typologically very distant from each other, such as Swedish, Italian and Japanese (Di Biase & Kawaguchi 2002, Kawaguchi 2011 and Glahn et al. 2001).

According to Pienemann, there are five of these procedures that build upon each other:

1. The lemma
2. The category procedure (lexical category of the lemma),
3. The phrasal procedure (instigated by the category of the head),
4. The S[entence]-procedure and the target language word order rules,
5. The subordinate clause procedure (if applicable).

(Pienemann 2011:33)

Pienemann asserts that there is an implicational relationship between these procedures: Each procedure is propaedeutical to the next, meaning, for instance, that the access to the lemma is a prerequisite to the category (morphology of the lemma) procedure. This, in turn, is a prerequisite to the phrasal procedure and so on. For instance, in order to be able to apply the plural form to a noun, the noun's morphological properties must be known: Properties such as gender and the rules for plural formation must be present in the learner's knowledge. Stage 1, the lemma, must be acquired before the learner is able to assign word classes to these lemmas (stage 2). When this happens, the learner will be able to use different plural endings. Consequently, in order to apply an adjective to a nominal phrase, both the properties of the adjective and noun (and therefore the grammatical rules of attributive congruence) must be processed by the learner. Without this procedure, the speaker can place an adjective before a noun, but, in the mind of the learner, they will not unite as a grammatical phrase and, thus, no grammatical information can be transferred between these two words. So, for a learner of L2 Swedish, the utterance *fina hundar* requires the acquisition of the first three procedures, namely lemma (*hund*, *fin*), category (*hund-ar*) and phrasal (*fin-a hund-ar*) procedures (Abrahamsson 2009:124). Once the learner can automatically (or "without thinking"

formally about it) use a procedure, he or she is ready to process the next one in the scale. Being able to perform this without using metalinguistic knowledge (formal thinking), indicates that the phrasal procedure is there, and it is, in theory, possible to develop to the next stage.

PT is based on the first emergence of a new grammatical structure, which, in the context of PT, means that as soon as a structure is used in a non-root manner or not coincidentally, it counts as acquired. Secondly, there is the matter of accuracy, where errors can, in fact, be an indicator of progress: When a speaker says *hund-er* instead of *hund-ar*, the grammatical category *plural* counts as acquired according to PT, even if the noun is not lemmatized. The error shows that the learner can process plural and is situated on PT level 2, although *hunder* is incorrect. It could be, in fact, even better evidence of acquisition, since it would show repetition and overgeneralization of a rule heard before. Pienemann argues that "the point of emergence is the only point in the acquisition process at which the two dimensions of Second Language Acquisition are not blended and that, in addition, it is also psychologically plausible because the emergence of a structure is the point at which the operation underlying the given structure is processable for the first time" (Keßler and Pienemann in Pienemann 2011:94). In other words, incorrect usage of a grammatical feature can be an indicator of development. It is important to note that PT is not a theory about acquiring plural, gender, agreement or word order, but it is about the natural development of language: The structures that we look at are simply indicators that the procedures are in place. A learner of a language without phrasal agreement must also acquire the phrasal procedure to be able to pass on to the S-procedure, but for this language we would have no way to see when this phrasal level is acquired.

The key thought behind PT is that second language development is closely linked to the grammatical processes and procedures which are responsible for common and spontaneous oral production. Language development occurs when these procedures are automatized in the learner's mind, which results into an expansion of his processing capabilities. This takes a long time and continues after the first emergence of processing. Structures are continuously automatized long after the procedural capacity is in place. In order to achieve a higher developmental stage, which requires

a higher processability capacity, the capacity for the current level must be reached first (Abrahamsson 2009:124).

PT identifies the grammatical exchange of information between constituents, the agreement between constituent parts of a given grammatical structure, as "one of the key factors upon which the processing and acquisition of many L2 structures depend". In the beginning, the learner cannot produce many L2 structures due to the constraints "imposed on L2 production by the learner's complete inability to deposit grammatical information into syntactic procedures" (Pienemann 2011:34).

2.2   Developmental stages for Swedish

Håkansson (Pienemann & Håkansson 1999) applied PT to Swedish and exemplified and adapted the five procedural levels into the Swedish language. The levels are described as follows:

Table 1. Processability levels in Swedish (adapted from Håkansson 2004:154, translated from Swedish by me)

Level	Procedure	Exemple from Swedish morphology	Exemple from Swedish syntax
Level 5	Grammatical information between clauses. Difference between main and subordinate clause		Negation after the verb in the main clause and before the verb in the subordinate clause
Level 4	Grammatical information between phrases	Predicative congruence	Normal and inverted word order
Level 3	Grammatical information within the phrase	Attributive congruence	Initial placement of adverbial modifier/object and normal word order
Level 2	Word class, lexical morphology	Suffixes for plural, present, preteritum	Canonical word order
Level 1	Individual words	Invariant forms	Individual constituents

At the first level, the learner can only process single words. Lemmas are here processed as single units, without any morphological aspect to them even though he or she may be able to form meaningful word clusters. At this level, there is no grammatical structure and no grammatical exchange.

In the second level, the learner can process word classes and assign them to lemmas. This enables him or her to apply the plural to a substantive (*hund-ar*) or conjugate a verb, maybe overutilizing, for instance, the most common plural suffixes (*-ar* and *-er*) even for nouns that do not require them. He can also understand that there are different word classes which have different functions. The word order is canonical, meaning the subject is always the first word in a sentence and the verb will always follow.

At the third level, the morphology is developed to function within the phrase. The learner is now able to process whole phrases and decline their constituents accordingly. The previously mentioned example *fin-a hund-ar* comes to mind. He is therefore able to transmit grammatical information within the same nominal phrase, since both the adjective and the noun are declined according to the Swedish plural rules.

On level four, grammatical information between phrases can be processed and the learner can start utilizing predicative congruence. Thus, the exemple *hund-ar-na är fin-a*, in which the noun phrase and the predicative phrase are in accordance: The adjective *fin* is declined in plural after the noun it refers to. For the syntax, the learner is now able to use both canonical and inverted word order. This means that, after placing an object or an adverbial in the beginning of the sentence, the learner can now place the verb after the fundament and invert the word order, complying with the Swedish syntactical rules.

Lastly, in the highest level of the PT hierarchy, the learner can process the exchange of grammatical information not only between words and phrases but also between clauses. This means, in Swedish, that the learner can distinguish between main and subordinate clauses and can apply syntactical rules such the placement of the negation in the subordinate clause and the inversion of verb and subject in the

main clause if the fundament is constituted by a subordinate clause (Håkansson 2004:153-156).

For the purpose of this paper, it is necessary to furtherly examine the developmental sequence for the adjective. In Swedish, the adjective agrees with the noun in number, gender and definiteness. The distinction between the two genders, uter and neutrum, is expressed respectively with the base form of the adjective and the use of the suffix -*t* in the indefinite form. The definite form will not be used in the present experimental study because the adjective agrees with definiteness only within the noun phrase, and not in predicative position. It would therefore be impossible to make a comparison for that type of agreement. The number function is instead identified by adding the suffix vowel -*a* to the adjective. There is therefore no gender expression in the plural. Agreements applies to both attributive and predicative position, so in both noun and verb phrase (Glahn et al. 2001:393).

Glahn et al. (2001) engaged in a study aimed to establish and verify the learning order of the adjective (in both attributive and predicative position) and, moreover, the acquisition of the negation in the subordinate clause. The research, therefore, covered the levels 3 to 5 of PT. Similarly, to the present work, the researchers used pictures in order to elicit the usage of color adjectives in the 47 participants to the study. This study, thanks to the extensive use of implicational tables for the different grammatical phenomena analyzed, suggested that "phrasal adjective agreement is used more frequently than interphrasal adjective agreement, which in turn is used more frequently than subordinate-clause negation placement" (Glahn et al. 2001:400). Glahn et al (2001) also proved Pienemann's acquisition schedule: "Phrasal morphology is acquired before interphrasal morphology, which again is acquired before the preverbial position of the negation in the subordinate clauses" (p. 413). And, as mentioned earlier, the study showed that there is a number-gender hierarchy that parallels the phrasal-interphrasal hierarchy, meaning that the acquisition of these morphological categories followed a fixed order similarly to ATTR-PRED and that the acquisition of the number precedes the acquisition of the gender.

As for the developmental sequence, it is exemplified in the following table:

Table 2. Developmental stages of the adjective

Stage	Exemple
1. Incongruence (non-inflected adjective)	*Jag älskar röd jordgubbar *Jordgubbar är röd
2. Agreement in attributive position	Jag älskar röda jordgubbar *Jordgubbar är röd
3. Agreement in predicative position	Jordgubbar är röda

At the first stage, there are only uninflected adjectives, meaning that agreement with the noun has not manifested itself in the learner's production. At the second stage, the learner processes agreement between noun and adjective in the attributive position, meaning adjective in the nominal phrase. At the third level, the learner can inflect the adjective in the predicative position after the noun it refers to, thus letting two different phrase "communicate" with each other. This is also in agreement with the stages 2 to 4 of PT, where the exchange of information goes from within the same word to within the same phrase to between two or more phrases (Abrahamsson 2009:75-76).

2.3 Assessing second sanguage proficiency: The Common European Framework of Reference for Languages

CEFR provides, in the present study, the stages used to assess the experimental texts. The Common European Framework of Reference for Languages

"…provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe. It describes in a comprehensive way what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively. The description also covers the cultural context in which language is set. The Framework also defines levels of proficiency which allow learners' progress to be measured at each stage of learning and on a life-long basis."

(Council of Europe 2001:1)

CEFR was published in 2001 and it is the result of a standardizing work that has been continuous since the 1970's. The framework sees to the learner as a "social

agent", which means an individual living in a society with other individual who must be able to perform certain tasks in their everyday life. These tasks require communicating with other people, therefore CEFR is said to be communication oriented. A significative feature of CEFR is that it takes into consideration "the cognitive, emotional and volitional resources and the full range of abilities specific to and applied by the individual as a social agent" (Council of Europe 2001:9).

The document describes the learner's competences and divides them between general and communicative language competences. The communicative language competences, useful in the context of the present paper, are divided in linguistic, sociolinguistic and pragmatic competences. The linguistic competences are furtherly divided into lexical, grammatical, semantic, phonological orthographical and orthoepic. For each of these competences, CEFR established an assessment table based on what the user can do, and which describes the user's competence for each skill in detail. CEFR is perhaps most known for its proficiency levels, from A1 to C2, which can be used for every language represented in the Council of Europe.

While it has been praised for providing a common framework for assessing European language proficiency, CEFR has also been critiqued for several reasons. Kuiken and Vedder (2014) cite the "rather general, descriptive nature of the scales" (2014:282) as problematic because of its vagueness and interpretability, which is of course of individual nature and varies with every assessor. Hulstijn (2007), instead, focused on the lack of empirical support for what being on a certain level means for a learner, in terms of performance. Moreover, it is possible for a learner to find him or herself on non-matching levels of several proficiency scales, which could be problematic when assessing on which level on the overall scale the learner is.

Despite these critiques, CEFR remains one of the most used tools for assessing language proficiency. The Common Reference Levels are depicted in table 3 and they start from A1 (Breakthrough, lowest) until reaching C2 (Mastery, highest) which represents a native speaker's proficiency level.

Table 3. Common Reference Levels

A		B		C	
Basic User		Independent User		Proficient User	
A1	A2	B1	B2	C1	C2
Breakthrough	Waystage	Threshold	Vantage	Effective Operational Proficiency	Mastery

(Council of Europe 2001:23)

For the purpose of this paper, the only scale that will be reported here is the one relative to the grammatical competence, which will be later use in the assessment of the learner texts. This competence is defined in CEFR as "knowledge of, and ability to use, the grammatical resources of a language (2001:112). Table 4 is one of the many competence-specific scales found in CEFR, and it focuses on grammatical accuracy. To each CEFR level (A1 through C2) corresponds a description of the learner's ability to perform in the target language.

Table 4. CEFR Scale and Descriptors for Grammatical Accuracy

	Grammatical Accuracy
C2	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).
C1	Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot
B2	Good grammatical control; occasional 'slips' or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.
	Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding.
B1	Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influence. Errors occur, but it is clear what he/she is trying to express.
	Uses reasonably accurately a repertoire of frequently used 'routines' and patterns associated with more predictable situations.
A2	Uses some simple structures correctly, but still systematically makes basic mistakes – for example tends to mix up tenses and forget to mark agreement; nevertheless, it is usually clear what he/she is trying to say.
A1	Shows only limited control of a few simple grammatical structures and sentence patterns in a learnt repertoire

(Council of Europe 2001:114)

2.4   Previous research on the relation between PT and proficiency

Several studies have investigated the relationship between second language development (L2D) and second language proficiency (L2P), often using PT and developmental sequences as main framework of L2D.

Granfeldt & Ågren (2013) assessed, in a small-scale pilot study, 38 learner texts of Swedish students of French according to CEFR (Common European Framework of Reference for Languages) guidelines, using them as a method of measuring L2P. The same texts were then analyzed following PT. The researchers found strong evidence of a linear correlation between CEFR ratings and PT-levels at lower levels of proficiency (specifically A2 and A1), meaning that a higher proficiency level generally corresponds to a higher PT level. They also discovered the presence of "uneven profiles in the data, i.e. learners with stronger communicative proficiency than morphosyntactic development or vice versa", which becomes more frequent at more advanced stages (2013:37).

A comprehensive study investigating the relationship between PT and L2P was done by Eklund Heinonen (2009). This study analyzed TISUS (Test in Swedish for University Studies) takers and their level of second language development according to the processability levels. In the study, it is argued that the *emergence criterion*, also known as *first occurrence criterion*, is too unreliable to be used in an assessment setting, since most of the test takers were relatively advanced users of Swedish. However, by using the more stringent criteria for analysis of 50% and 80% of obligatory contexts, a clear difference was demonstrated between the groups who passed and the group who did not. The result of the study was that most of the applicants who passed the exam were at PT4 or higher, whereas most of those who failed were at PT3. Therefore, the study found a correlation between proficiency (TISUS results) and second language development (in terms of PT levels).

3. Method

This chapter is going to describe the experimental method used in this study. Firstly, the experimental method used in the present study will be introduced. Secondly, an account of how the method was applied will be provided, with a description of the process of text creation and manipulation. Thirdly, a paragraph describing the participants will follow. The final part of this section will instead explain how the result will be analyzed.

3.1   The experimental method

One of the positive aspects of choosing an experimental method for a linguistics study is the possibility of controlling the variables that will be introduced into the study. The main difference with other methods used in linguistics is that the experiment is more theory and hypothesis oriented, meaning that a clear hypothesis must be established before proceeding with designing the experimental study. According to Clark & Clark (Cited in Lagerholm 2010:59), the design of an experimental study consists of several phases, described as follows:

1. Theory: The researcher has a theory that must be tested in some way.
2. Prediction: Starting from the theory, the researcher derives a prediction in the form of if *X occurs then Y must occur.*
3. Manipulation: In the typical experiment, the conditions are manipulated so that X occurs.
4. Observation: The researcher observes whether Y occurs.
5. Conclusion:
   5.1 If Y occurs, the prediction is therefore strengthened, and the researcher draws the conclusion that it supports the theory.

5.2 If Y does not occur, the researcher draws the conclusion that they have proof that goes against the theory – that the theory is partly or completely wrong.

5.3 If the observation is uncertain, the researcher does not draw any conclusion at all.

As the reader will have noticed, the experimental design can be a quite complex process and requires thorough knowledge of the primary theory used, in this case Pienemann's Processability Theory. Although the whole process is complex, there is an overwhelmingly positive aspect to this experimental design: After the study has been done, it is quite easy to compare the results, both within the study and between several studies, since the conditions should be easily reproduced by using the same the same variables.

There are negative factors too. There is always the risk that the experiment will be more difficult to both interpret and carry through than the researcher at first imagined. Moreover, the difficulty of controlling as many factors as possible is also to be considered when using this method. There is always the risk of not recognizing some factors as decisive to the study and, if not ignoring them, at least not giving them enough weight in the study. Another risk with the experimental method is that the situation analyzed is so far removed from a real-world situation that the results will be inapplicable.

After choosing the Processability Theory as the theory to be used in this study, the researcher formulated a prediction, or hypothesis. We know that, based on numerous studies (see Chapter 2), there is a link between proficiency and PT. What the present study will instead analyze is whether this relationship is causal. In other words, if the relationship is causal, the assessment of learner texts will (consciously or unconsciously) be influenced by the PT level. If, instead, the relation is not causal, the assessors will not pay attention to the PT level and the texts with higher PT level would simply tend to be better in other respects.

3.2   Material

In order to ascertain if the PT level does indeed affect grading, four texts have been manipulated so that the only element that differentiates them is the PT level, while at the same time controlling every other factor. The texts were then submitted to teachers of Swedish as a second language for grading and assessment. If the PT level has an influence over the grade, it will mean that there will be a difference in grade assessment between two versions of the same text. If, instead, there is no significative difference in grading between the two version, the relationship between PT and proficiency is not likely to be causal.

In order to keep the experiment as pure as possible, I have elected to include exclusively adjectival errors in both predicative and attributive position. Adding more types of errors would have increased the unreliability of the experiment, since it would have made more difficult to decide whether the grading would be influenced by the PT level or by the other types of errors. Adding, for instance, syntactical errors, would have an impact on the grading that would go beyond the analysis of PT levels. With this design, no other PT related errors could be tested. If, for instance, we decided to test word order, we would have to compare a correct word order with an erroneous word order. In that case, we could not know whether the difference would be because of the error per se or because of the PT level. Thus, adjectives are perfect for this study, since the "same" error would yield a different PT analysis depending on whether it is in attributive or predicative position.

The texts were created by describing a series of four pictures depicting urban sceneries and people in the middle of different acts. This kind of picture describing process is often used as a second language learning exercise at a basic level. This type of exercise was chosen in order to keep both the texts and the survey as short and simple as possible. Therefore, each text contains fewer than one hundred words. It was decided that, in order to keep them simple enough, they would not contain any other type of PT-related structure that was not the adjective in predicative and attributive position. There are no other structures that could indicate that the text could be at level 4 or 5 (other than the adjectives in predicative position).

Each text was made to contain in total ten adjectival contexts. The reasoning behind this choice was to create two versions of the texts: The PT3 version, assessed to be PT level 3 according to standard measures (see background part of the present paper), and the PT4 version, located at PT level 4 according to the same set of rules. Incorrect structures both related and unrelated to PT were also not included in the manipulated texts. If, for instance, the PT3 version had examples of incorrect word order while at the same time the PT4 version had correct word order, it would be impossible to say whether a difference in assessed proficiency would be due to PT level or general accuracy. The PT3 versions contain five correct instances of the adjective in the attributive position and zero correct instances of the adjective in the predicative position, whereas the PT4 versions contain three correct instances of the adjective in the attributive position and two correct instances of the adjective in the predicative position. As explained in the theory section, two cases of agreement should be enough to claim that the learner has reached level 4 of PT.

Importantly, the number of errors is the same in the two versions, to ensure that general accuracy would not lead to a difference in assessment. Each version contains, therefore, 5 errors in 10 total adjectival contexts. The errors in the two texts are also very much "the same", except for some ungrammatical adjectives in attributive position, while others are in predicative position. For instance, Text 1 variants 1-PT3 and 1-PT4 (see Appendix) contain five incorrect uses of the adjectives; for Text 1-PT3 all of them are predicative contexts, while in 1-PT4 they're also five, but distributed as two attributive and three predicative contexts. So, the correct *en stor brand* in 1-PT3 becomes *ett stor brand in 1-PT4, while **tre personer är nyfiken* in 1-PT3 becomes the correct tre *personer är nyfikna* in 1-PT4. It was, moreover, ensured that the errors in both versions are the same regarding what their function is in stead of the form (gender, definiteness and number). A few filler errors were also included in the texts, which will be described in a separate section.

The texts were then submitted to teachers of Swedish as a second language for assessment. First, I conducted a small-scale pilot study to eight teachers belonging to several schools in the Gothenburg area, who were asked to assess the texts. This was done in order to verify that the texts were balanced, and that no assessment would be too one-sided. The result of the pilot study will, therefore, not be reported.

The texts variations were then split into two lists, each one containing four texts, and submitted to the assessors via online survey. Each list contains two PT3 texts and two PT4 texts (see table 5).

Table 5. Lists and texts distribution

Text	List 1	List 2
1	1-PT3	1-PT4
2	2-PT4	2-PT3
3	3-PT3	3-PT4
4	4-PT4	4-PT3

The texts can be found in the Appendix. The reader will notice an abundance of ett-nouns in every text, which does not mirror the percentage of ett-nouns in the Swedish language. This was decided because the novice Swedish learner tends to use the default form of the adjective when applied to neutrum and plural nouns, ignoring any morphological flexion of the same. It would be therefore unnatural to find adjectives flexed in the neutrum form referring to non-neutrum nouns. For instance, *sanden är brunt* would be an unnatural construction. Thus, to make it very clear whether the imagined learner had acquired adjectival agreement in attributive and predicative position, we had to use nouns belonging to neuter gender and plurals. Consequently, plural forms were also used in abundance for the same reason. It was more natural to turn the correct *några trevliga killar* into the incorrect *några trevlig killar* in the relative text for the purpose of this research rather than using the singular *en trevligt* (or *en trevliga*) *kille*.

3.2.1 Assessment scale

In order to grade the paper, I created a 6-graded matrix based on the CEFR scale for grammatical accuracy (see Section 2), where each definition would correspond to a grade from F (lowest) to A (highest). I chose this scale among all the scales contained in CEFR because it focuses on grammatical competence and it is extremely relevant to the object of this study. A positive aspect of a grading scale is its familiarity for the assessors: Many teachers already use this kind of matrix for their work, often based on the official language proficiency criteria given out by the Swedish National

Agency for Education (which, in turn, are in great part inspired by CEFR). Since the experiment was conducted via online forms, the room for the criteria used was limited by practical constraints. Firstly, the online tool used did not allow for many characters to be used. Secondly, by using the full definitions of CEFR criteria there was the risk of confusing the assessors, with the possible risk of them abandoning the survey entirely. Therefore, I chose to shorten the definitions given in Table 4 in order to fit the format of the survey. The modified scale with definitions is taken from the Swedish translation of CEFR. Scale and matching grades are reported in table 6.

Table 6. Grades, CEFR levels and shortened definitions

Grade	CEFR	Definition
F	A1	Eleven visar endast begränsad kontroll över några få enkla grammatiska strukturer och satsmönster inom en inlärd repertoar.
E	A2	Eleven använder några enkla strukturer korrekt, men gör fortfarande systematiskt grundläggande fel. Det framgår vanligen tydligt vad han/hon försöker säga.
D	B1	Eleven kommunicerar någorlunda korrekt i välbekanta sammanhang och har i regel god kontroll. Fel förekommer, men det framgår tydligt vad han/hon försöker säga.
C	B2	Eleven har god grammatisk behärskning, även om det fortfarande förekommer enstaka felsägningar och icke-systematiska fel och små misstag i meningsstrukturen, vilka dock är sällsynta och ofta kan korrigeras i efterhand.
B	C1	Eleven visar genomgående en hög grad av grammatisk korrekthet; gör sällan fel och eventuella fel är svåra att upptäcka.
A	C2	A: Eleven visar genomgående grammatisk behärskning över komplext språk, även när uppmärksamheten är riktad på annat.

3.2.2 Fillers

Filler errors were included into every text, so that the assessors would not "suspect" the real scope of this investigation. They are intended to work as diversion and to "hide" the objects of the experiment (the adjectival contexts) in a layer of more superficial errors. All the fillers are orthographic and lexical errors which do not hinder communication. So, for instance, in Text 1 we have **brändman* and **brändbilar*, instead of *brandman* and *brandbilar*. A person with Swedish as L1 would still understand what the learner meant, inferring from both the context and

their own lexical knowledge. At the same time, we have *å, which takes the place of the conjunction *och*, as they are pronounced the same way. Even here, communication is not problematic, and the reader can easily see what the learner intended to write. A List of fillers in table form can be found in the Appendix.

3.3 Participants

In total, 62 assessors took part to the present study. The surveys were shared on the internet forum known as *Reddit*, in the "subreddit" r/Svenska. They were also shared on the following Facebook groups dedicated to Swedish as a second language: *Svenska som andraspråk*, *Studera svensk grammatik* and *Vi studerar Svenska*. There is in practice no difference between the groups: The only requirement asked in order to participate to the survey was to be either a teacher of Swedish as a second language or a student of the same subject.

The survey asked all the participants to explain for how long they had been teaching Swedish, although it was an optional question. The level of individual experience varied from 0 (students or new teachers) to 30 years. The mean experience for the List 1 group is 5.23 years, while for the List 2 group it is 6.53 years. It is, therefore, very similar for both groups and does not relate to the mean grade at all. Two assessors were excluded from the study, because the average of their grades was considered extremely low or extremely high and could be considered outliers.

3.4 Data analysis

A numeric value was then assigned to each grade, which ranged from 1 (lowest) to 6 (highest) (see Table 7). This was done in order to make calculated average values possible for each text. The mean grades were then calculated and compared to establish possible patterns in the results. The averages were calculated separately for each group and they include mean individual grade, mean experience and mean grade for each text. The mean grades of the PT3 and PT4 versions of the same text were

then compared to each other, in order to establish if there was a significant difference in grading.

Table 7. CEFR levels, grades and their assigned numeric values

CEFR Grammatical Accuracy	Grade	Numeric Value
A1	F (lowest)	1
A2	E	2
B1	D	3
B2	C	4
C1	B	5
C2	A (highest)	6

4.   Result

This section will show the results of the experimental study described in Chapter 3. The average grade for each version of the texts will be calculated and displayed, followed by a direct comparison of the two versions of the same text.

4.1   Analysis of the result

A full account of the result can be found in the Appendix section. The next table will instead summarize the spread (minimum and maximum numeric value) for each text:

Table 8. Spread between highest and lowest grades for each text

Text	Highest Grade	Lowest Grade	Spread
1-PT3	4	1	3
1-PT4	4	1	3
2-PT3	5	2	3
2-PT4	4	1	3
3-PT3	5	1	4
3-PT4	4	1	3
4-PT3	4	1	3
4-PT4	4	1	3

By looking at these values, it would appear that the participants are in general agreement on a range of grades. They have settled on a grading range that goes from 1 (F) to 5 (B); none of the participants gave any of the texts the highest possible grade, 6 (A).

In order to compare the PT3 and PT4 texts, I have compiled a table using the mean grade given to each text. Two participants were excluded by list 2 because their

grades diverged from the mean by more than 2 standard deviations and can be considered outliers.

Table 9. Mean grades for each text and spread

| PT3 Text	Mean Grade	PT4 Text	Mean Grade	PT4 – PT3 Difference
1-PT3	2.68	1-PT4	2.94	+0.26
2-PT3	3.23	2-PT4	3.06	-0.17
3-PT3	3.31	3-PT4	2.94	-0.37
4-PT3	2.77	4-PT4	2.50	-0.27

We see in the table that there are only slight differences in the assessment of the PT3 and PT4 versions. It appears, therefore, to be a minimal spread, in the range ±0.37. By examining the results of the investigation for Text 1, it was shown that the PT4 version received a better mean grade than the PT3 version. The PT4 versions of the other texts did, however, receive a better grade than the PT3 variants, although in both case the difference in grade between the versions is basically zero. Where the Text 1-PT3 grade was higher than Text 1-PT4's by 0.26 points, the differences in the other texts was therefore, instead, negative: Text 2-PT4 was assessed to be better than text 2-PT3 by 0.17 points, Text 3-PT4 better than 3-PT3 by 0.37 points and Text 4-PT4 better than 4-PT3 by 0.27 points. It is important to stress that since the differences amount at maximum to about a fourth of a grade, they are, for all purposes and intents, irrelevant: There is no actual spread in the grading values.

4.2   Assessment and individual experience

To investigate a possible relationship between experience in teaching Swedish as a second language and grades assigned to the text, a graph was drawn, illustrating the distribution of grades per years of experience. This was done in order to be sure that no irrelevant factor affected the result. One of these factors could have been a possible disparity in teaching experience between the list groups: If, for instance, one group was in average more experienced than the other, this could have affected the result. It is shown instead that the average experience for the first group is 5.23 years, while for the second group it is 6.53 years. They can be, therefore, considered equal.

A graph was then drawn to emphasize any possible relation between experience and grade. Teacher students, teachers with less than a year of experience and teachers who did not submit how much experience they had, were counted as having 0 years of experience.

Table 10. Grades distribution for years of experience
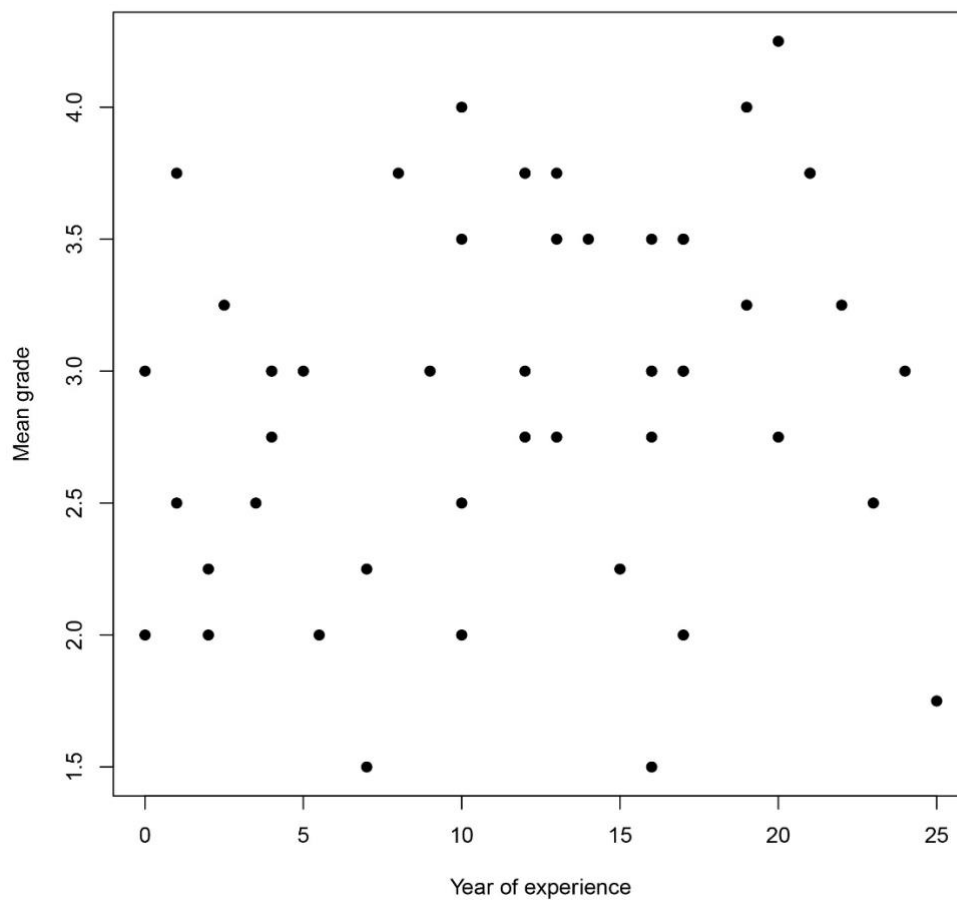
Each dot represents a teacher. The distribution of the dots shows no correlation whatsoever between the variables Mean Grade and Years of Experience.

5. Discussion

5.1 Summary of the study

Previous studies have found correlations between level of processability, as measured with the Processability Theory, and general language proficiency. This experimental study has tested whether such correlations are indicative of a causal correlation or are rather the effect of an indirect relation between proficiency and level of processability. In an online survey, 63 Swedish teachers rated four texts for grammar according to the CEFR. The texts were manipulated so that they were either on PT level 3 or 4. By focusing on adjectival agreement, the texts could be manipulated so that the number of mistakes and also the "nature" of mistakes where the same in the two versions of each text; the only difference was that in one version, the texts showed evidence of processing only attribute adjectives (indicative for level 3) whereas in the other version, the texts showed evidence of processing both attribute and predicative adjectives (indicative of level 4). There was no difference in grading depending on the PT level. In 3 cases out of 4, there was instead a very small difference in favor of the PT3 version of the text. By translating the numeric values assigned to the rating into our original grading system, the mean grades would range roughly from D- to D+. The results clearly show that the teachers were not sensitive to the difference between PT levels: There was no relation between level of processability and assessed level of proficiency. This suggests that the correlation between PT level and proficiency found in previous studies is indicative of an indirect relation: As speakers become more proficient, their grammatical processing capacity also develop, but assessors do not pay attention to this development specifically, even when focusing on grammar, as in this experiment.

5.2 Methodological discussion

An aspect that may have influenced the outcome of the investigation is the identity of the assessors. Since the investigation was conducted via anonymous online

polling, it was impossible to determine, or rather, certify, whether the assessors would really be teachers of Swedish as a second language or just unrelated individuals, which is also due to the nature of the present paper and resource constraints that applied to me as a university student. The ideal profiles of the assessors would all have the same formal training in Swedish as a second language, in order to reach the most accurate possible outcome in the research. With each assessor sharing the same knowledge, the texts would have been graded under much more similar circumstances, reducing the individual difference between the assessors and minimizing the risk of impersonation by people unrelated to the task. A difference in experience or "generosity" with the ratings could also have affected the result of the study. However, this was not shown not to be the case: The difference in grades between the groups related to each list is minimal and the averages are very much the same. Even the reported average experience is the same and, moreover, it has been shown that experience and grade given do not correlate. It should also be pointed out that the participants were recruited in Facebook groups for Swedish as L2.

Another methodological issue that I have encountered is that, being the present study based on experimental texts created ad hoc, the assessors sometimes recognized the texts as not authentic. I received two private messages on the platforms where the polls were published asking me whether the texts were written by learners. This led me to think that at least some of the assessors recognized them as non authentic, albeit none of them openly referred to the Processability Theory or even any other second language acquisition topic. However, the suspicion of working with a non-authentic text might have put the assessors in a different mindset compared to what they would have by working with texts created by students of Swedish as a second language. Despite the very careful work put into text creation by me, perfectly simulating authentic texts would have been a nearly impossible task, given the constraints required by the experimental method and the research question. In any case, even if some assessors recognized the texts as non-authentic, they did not perceive how the texts were manipulated, and even if they did, it would not affect the result in any imaginable way. Thus, the internal validity of the experiment is high, since PT has shown no effect on the assessors.

5.3   Discussion of the result and previous research

Agebjörn (2019) analyzed the assessment of twenty learner texts created by students of Swedish, with a focus on grammatical complexity and accuracy. The results of the study showed that, in general, assessors are influenced more by the grammatical accuracy, defined as the least possible amount of errors, in a given text rather than the complexity, here exemplified by the PT scale and related to the amount of grammatical structures that a learner is able to use, despite the amount of errors. This is in line with the results of this study: Given two equal texts that only differ between each other in the type of error while keeping the same amount of errors, the assessor will not focus on the grammatical complexity of the text (as defined by PT, that is, what grammatical processes take place in the learner).

Gyllstad et al. (2013) used instead three criteria to measure L2 complexity in 54 learners of English and 38 learners of French. These criteria are based on general measures for complexity, such as length of sentences, clauses and phrases. The results suggest that complexity measured through these criteria does not automatically translate into better communicative proficiency. This is also in line with the results of the present study, where texts placed at different PT levels received almost identical grades. It is not clear what the assessors are focusing on: It could be accuracy or other aspects of the texts; but the present study indicates that they are not paying attention to the PT level.

Even though we found no correlation between PT and grade, that does not mean that PT cannot be relevant in assessing a text. It might be the case that, as suggested in Agebjörn (2019), assessors focus on accuracy, leaving behind the complexity. Complexity could be codified by the PT scale. The CEFR scale used in the present study is, in this case, vague, as it mentions grammatical accuracy, whereas linguistic complexity is only mentioned at the top of the scale, level C2 (Council of Europe 2001:114), and is never defined, leaving to the assessor the task to interpret what *complex language* means. Gyllstad et al. (2013:104) found that the notion of complexity is present in 149 total cases in the whole document and in four out of six levels in two scale taken into exam (Gyllstad et al. 2013:105) but, despite this, the document shows no adequate definition for complexity. PT could, possibly, provide

help. CEFR definitions such as "simple grammatical structures and sentence patterns" and "complex language" present the problem of being vague; PT would, instead, substitute them with very specific descriptors of what the learner is able to process. While the scope of the present study is very specific and it only relates to the causality between PT and proficiency, it could help further research that, for instance, could investigate whether assessment reliability would increase with CEFR levels that define complexity with PT.

A critique to CEFR is that it is focused on the communicative approach to language learning, while ignoring the specific linguistic competence of the learner and the fact that "pragmatic and conversational competence are realized primarily by means of linguistic resources […] What language learners can do with language is to a very considerable extent dependent on what language they know" (Ellis 2008:17-18). Moreover, whereas CEFR sees at errors as something that might hinder communication, PT looks at them as signs of linguistic development in the learner. My point is that by codifying PT into a complexity scale, it would be possible to make the grading process more reliable, thus certifying that, despite the errors present in a text, the learner is ready or is already using a more advanced grammatical structure (such as the adjective in predicative position). Although PT can, by its own nature, focus only on a small part of the learner's linguistic competence, there could be place for it in the grading system. An immediate practical use of this new scale would be in grading student texts, to let the learner know that he or she is able to produce certain grammatical structures, even despite the eventual errors; and that information would make explaining the errors to the learner an easier task for the teacher.

5.4   Further research

It would be interesting to repeat my experiment with two groups of assessors. One group would have received formal training in PT, while the other would not. With this, it would be possible to analyze whether the acquaintance with the Processability Theory would make a difference in the evaluation of the texts. Would it be likely for the assessors trained in the use of PT would be affected by the theory and, possibly,

would assign higher grades to the PT4 versions, since they would show a higher grade of grammatical complexity (according to PT) than the PT3 texts? It would also be interesting to see if and how they would be consciously or unconsciously more sensitive to the PT difference between the two versions. A smaller scale study could even ask them to report their reasoning behind every grade, where they would have the chance to explain whether PT has affected the process of grading at all, or if they have taken advantage of any other kind of second language acquisition theory. In other words, they would be explaining in detail the reasoning behind every grade. The group without formal training in PT would also be asked to report their reasoning, since it would be very interesting to know also their own definitions for each grade.

6.   Conclusion

The Processability Theory is a second language acquisition model that might have an impact in how we look at learner texts and how they are assessed. In order to ascertain this, an experimental study has been conducted, where texts were manipulated into PT3 and PT4 versions. The grades given by the assessing teachers showed that the PT level indeed does not affect the grading of the text, since they do not vary significantly. However, it is argued that PT could play a role in language proficiency assessment. This role is exemplified by more accurate descriptors for language complexity in grading matrixes that PT could provide.

7.   References

Abrahamsson, Niclas 2009. *Andraspråksinlärning*. Lund: Studentlitteratur.

Agebjörn, Anders 2019. Grammatical accuracy and complexity in a speaking proficiency test. I: Arntzen et al. (2019), pp.137-159. *Processability Approaches to Language Acquisition Research & Teaching*, Amsterdam: John Benjamins Publishing Company.

Council of Europe 2001. Common European Framework of Reference for Languages. <https://www.coe.int/en/web/common-european-framework-reference-languages>

Di Biase, B. & Kawaguchi, S., 2002. Exploring the typological plausibility of Processability Theory: language development in Italian second language and Japanese second language. *Second Language Research*, 18(3), pp.274–302.

Eklund Heinonen, M. Uppsala universitet. Institutionen för nordiska språk, 2009. *Processbarhet på prov: bedömning av muntlig språkfärdighet hos vuxna andraspråksinlärare*, Uppsala: Uppsala universitet: Swedish Science Press

Ellis, R., 2008. Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics*, 18(1), pp.4–22.

Glahn, Esther, Håkansson, Gisela, Hammarberg, Björn, Holmen, Anne, Hvenekilde, Anne, 2001. Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition*, 23(3), pp.389–416.

Granfeldt, Jonas, Ågren, Malin, Flyman-Mattsson, Anna & Norrby, Catrin 2013. Stages of Processability and Levels of Proficiency in the Common European Framework of Reference for Languages. The Case of L3 French. In pp. 28–38.

Gyllstad, Henrik, Granfeldt, Jonas & Källkvist, Marie, 2013. Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in L2 English and L3 French. In: Språk i undervisning: Rapport från ASLA:s vårsymposium, Linköping, 11-12 maj, 2012 pp. 103–113.

Hulstijn, J.H., 2007. The Shaky Ground Beneath the CEFR: Quantitative and Qualitative Dimensions of Language Proficiency 1. *Modern Language Journal*, 91(4), pp.663–667.

Håkansson, Gisela, 2004. Utveckling och variation – en tillämpning av processbarhetsteorin på svenskt inlärarspråk. I: Hyltenstam, K. & Lindberg, I., (2004), pp.151-167. *Svenska som andraspråk: i forskning, undervisning och samhälle*, Lund: Studentlitteratur.

Lagerholm, P., 2005. *Språkvetenskapliga uppsatser*, Lund: Studentlitteratur.

Kawaguchi, Satomi, 2011. Japanese as a second language: A test case for the typological plausibility of PT. I: Pienemann (2011), pp.99-105.

Keßler, Jörg-U. & Pienemann, M, 2011. Research methodology: How do we know about developmental schedules? I; Pienemann (2011), pp.84-96.

Kuiken, F. & Vedder, I., 2014. Raters' decisions, rating procedures and rating scales. *Language Testing*, 31(3), pp.279–284.

Pienemann, Manfred & Håkansson, Gisela, 1999. A unified approach towards the development of Swedish as L2: a processability account. *Studies In Second Language Acquisition*, 21(3), pp.383–420.

Pienemann, M. & Keßler, J.-U., 2011. *Studying Processability Theory: An Introductory Textbook*, Amsterdam: John Benjamins Publishing Company.

Selinker, L., 1972. Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1-4), pp. 209-232

8.   Appendix

The first part of this section will show every text variant used in the experiment. Adjectival contexts have been highlighted in tables right below the text they refer to. The second section shows, instead, the assessment of each text made by teachers of Swedish as a second language. Average values have also been calculated and are displayed in the tables.

Text 1-PT3

5/5 correct attributive contexts   0/5 correct predicative contexts

Ett hus är gul och brinner och ligger vid sidan av två bruna hus. Ett hus är röd å brändmannen har vita kläder på sig. Tre personer står vid ett fönster, tittar ut och är nyfiken. En grå katt sitter vid ett litet fönster och sover. Rök kommer ut ur ett fönster medan en brändman står och släcker ut en stor brand med vatten. Ett flygplan är stor och några brändmän som talar med mannen är lång.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
					Ett hus är gul Ett hus är röd Ett flygplan är stor	…personer är nyfiken …brändmän är lång
Correct	En	Ett	Plural	En	Ett	Plural
	En grå katt En stor brand	Ett litet fönster	Vita kläder Två bruna hus			

34

Text 1-PT4

3/5 correct attributive contexts   2/5 correct predicative contexts

Ett hus är gult och brinner och ligger vid sidan av två bruna hus. Ett hus är röd å brändmannen har vit kläder på sig. Tre personer står vid ett fönster, tittar ut och är nyfikna. En grå katt sitter vid ett litet fönster och sover. Rök kommer ut ur ett fönster medan en brändman står och släcker ut ett stor brand med vatten. Ett flygplan är stor och några brändmän som talar med mannen är lång.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
		Ett stor brand	Vit kläder		Ett hus är röd Ett flygplan är stor	…brändmän är lång
Correct	En	Ett	Plural	En	Ett	Plural
	En grå katt	Ett litet fönster	Två bruna hus		Ett hus är gult	Tre personer är nyfikna

Text 2-PT3

5/5 correct attributive contexts   0/5 correct predicative contexts

Mannen på cykeln har en röd tröja och väntar vid övergångsstället; hans bälte är brun. Det finns tre långa kvinnor vid busshållplatsen och de väntar på bussen. Ett flyplan är vit och grå och flyger i himlen. Det finns också ett tåg i backgrunden med röda dörrar. Några bilar sitter fast i traffiken; de är grön och två bussar på gatan är gul. Ett stort hus i backgrunden har ett grönt tak.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
					Hans bälte är brun Ett flyplan är vit Ett flyplan är grå	Några bilar är grön Två bussar är gul
Correct	En	Ett	Plural	En	Ett	Plural
	En röd tröja	Ett stort hus Ett grönt tak	Röda dörrar Tre långa kvinnor			

Text 2-PT4

3/5 correct attributive contexts   2/5 correct predicative contexts

Mannen på cykeln har en röd tröja och väntar vid övergångsstället; hans bälte är brun. Det finns tre lång kvinnor vid busshållplatsen och de väntar på bussen. Ett flyplan är vit och grå och flyger i himlen. Det finns också ett tåg i backgrunden med röda dörrar. Några bilar sitter fast i traffiken; de är grön och två bussar på gatan är gul. Ett stort hus i backgrunden har ett grön tak.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
		Ett grön tak	Tre lång kvinnor		Ett flyplan är vit Ett flyplan är grå	Två bussar är gul
Correct	En	Ett	Plural	En	Ett	Plural
	En röd tröja	Ett stort hus	Röda dörrar		Hans bälte är brun	Några bilar är gröna

Text 3-PT3

5/5 correct attributive contexts   0/5 correct predicative contexts

Det finns tre bruna hus med en massa fönster. Två gamla kvinnor står på en balkong och tittar på gatan. Man kan se en tjej: hennes namn är Sarah och hon tittar på tv i lägenheten. Tre bilar på gatan är gul och är taxi. En kille går på skateboard; skateboardet är röd och killen har en liten hjälm på sitt huvud. Några trevliga killar går på gatan. En av dem heter Simon. Hans öga är blodig och huvudet är stor. En tjej köper blommar; de är gul, medan en kille köper tidningen i kiosken.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
					Skateboardet är röd Huvudet är stor Hans öga är blodig	Tre bilar är gul De är gul
Correct	En	Ett	Plural	En	Ett	Plural
	En liten hjälm	Sitt huvud	Tre bruna hus Två gamla kvinnor Några trevliga killar			

38

Text 3-PT4

3/5 correct attributive contexts   2/5 correct predicative contexts

Det finns tre bruna hus med en massa fönster. Två gammal kvinnor står på en balkong och tittar på gatan. Man kan se en tjej: hennes namn är Sarah och hon tittar på tv i lägenheten. Tre bilar på gatan är gul och är taxi. En kille går på skateboard; skateboardet är rött och killen har en liten hjälm på sitt huvud. Några trevlig killar går på gatan. En av dem heter Simon. Hans öga är blodig och huvudet är stor. En tjej köper blommar; de är gula, medan en kille köper tidningen i kiosken.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
			Två gammal kvinnor Några trevlig killar		Huvudet är stor Hans öga är blodig	Tre bilar är gul
Correct	En	Ett	Plural	En	Ett	Plural
	En liten hjälm	Sitt huvud	Tre bruna hus		Skateboardet är rött	De är gula

Text 4-PT3

5/5 correct attributive contexts   0/5 correct predicative contexts

Det är många barn på en strand. Et skepp på havet är gul. Ett ungt barn liggar på en handduk och en tjej med brunt hår liggar på en röd handduk och solar. Tre ankor sitter på stranden; golvet är brun men ankorna är vit. Det finns ett rött hus; de säljer fisk och chips vid sidan av stranden. Det finns okså två lastbilar som säljer blommor. Ett blomma är färgrik och lastbilarna är grön

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
					Ett skepp är gul Golvet är brun Ett blomma är färgrik	Ankorna är vit Lastbilarna är grön
Correct	En	Ett	Plural	En	Ett	Plural
	Röd handduk	Ett ungt barn Brunt hår Ett rött hus	Många barn			

Text 4-PT4

3/5 correct attributive contexts   2/5 correct predicative contexts

Det är många barn på en strand. Et skepp på havet är gul. Et ung barn liggar på en handduk och en tjej med brunt hår liggar på en röd handduk och solar. Tre ankor sitter på stranden; golvet är brun men ankorna är vit. Det finns ett röd hus; de säljer fisk och chips vid sidan av stranden. Det finns okså två lastbilar som säljer blommor. En blomma är färgrik och lastbilarna är grön.

	Attributive Contexts			Predicative Contexts		
Incorrect	En	Ett	Plural	En	Ett	Plural
		Ett ung barn Ett röd hus			Golvet är brun	Ankorna är vit Lastbilarna är grön
Correct	En	Ett	Plural	En	Ett	Plural
	Röd handduk	Brunt hår	Många barn	En blomma är färgrik	Ett skepp är gult	

Fillers

Text 1		Text 2	
Correct	Incorrect	Correct	Incorrect
Brandman	Brändman	Flygplan	Flyplan
Och	Å	Bakgrunden	Backgrunden
Flammor	Flammar	Trafiken	Traffiken

Text 3		Text 4	
Correct	Incorrect	Correct	Incorrect
Fönster	Fonster	Ligger	Liggar
Se	See	Ett	Et
blommor	blommar	också	okså

List 1

Subject	Experience	1-PT3 Grade	2-PT4 Grade	3-PT3 Grade	4-PT4 Grade	Average
T1	10 years	D	D	E	D	2.75
T2	5 years	D	D	C	E	3
T3	5 years	E	D	C	D	3
T4	6 years	D	D	C	D	3.25
T5	5 years	D	D	D	D	3
T6	20 years	D	E	C	E	2.75
T7	-	D	D	D	C	3.25
T8	-	E	D	D	D	2.5
T9	5 years	D	C	C	D	3.5
T10	3.5 years	C	C	D	D	3.5
T11	-	C	E	D	F	2.5
T12	0 years	E	D	F	E	2
T13	15 years	F	E	E	F	1.5
T14	19 years	E	D	C	D	3
T15	Student	D	D	D	D	3
T16	5 years	C	D	D	E	3
T17	2 years	C	C	B	D	4
T18	some	F	D	F	E	1.75
T19	4 years	D	C	B	D	3.75
T20	12 years	E	D	C	D	3
T21	4 years	D	D	C	E	3
T22	-	E	E	D	E	2.25
T23	4 years	E	F	E	F	1.5
T24	6 years	C	C	C	C	4
T25	0 years	F	D	D	E	2
T26	-	E	D	D	E	2.5
T27	3 years	E	D	D	D	2.75

T28	-	E	D	D	E	2.5
T29	20 years	C	C	C	D	3.75
T30	5 years	D	C	B	E	3.5
T31	9 years	D	C	C	E	3.25
T32	-	D	D	D	D	3
Averages	5,23 years	2,68	3,06	3,31	2,50	2.89

List 2

Subject	Experience	1-PT4 Grade	2-PT3 Grade	3-PT4 Grade	4-PT3 Grade	Average
S1	10 years	D	D	D	D	3
S2	3 years	C	D	C	D	3.5
S3	-	D	D	E	E	2.5
S4	2 years	E	E	E	E	2
S5	2 years	C	D	C	D	3.5
S6	4 years	D	D	D	D	3
S7	5 years	C	C	D	D	3.5
S8	8 years	C	D	C	C	3.75
S9	30 years	E	D	E	E	2.25
S10	4 years	C	C	D	D	3.5
S11	18 years	C	C	C	D	3.75
S12	-	C	C	D	D	3.5
S13	5 years	E	D	E	F	2
S14	-	E	D	D	E	2.5
S15	20 years	D	D	D	D	3
S16	2 years	F	D	D	D	2.5
S17	Student	E	D	F	E	2
S18	-	E	D	D	C	3
S19	2.5 years	C	C	E	D	3.25
S20	5.5 years	E	E	E	E	2
S21	4 years	E	D	D	D	2.75
S22	7 years	B	B	C	D	4.25
S23	20 years	C	C	C	D	3.75
S24	7 years	D	C	E	E	2.75
S25	1 year	F	E	D	C	2.5
S26	-	E	D	D	F	2.25
S27	15 years	E	E	E	D	2.25
S28	3 years	D	C	C	C	3.75

S29	10 years	C	D	D	E	3
S30	8 years	D	C	C	C	3.75
Averages	6.53 years	2.94	3.23	2.94	2.77	2.97