Definiteness in written Swedish by Finnish-speaking immersion pupils at the end of immersion

A comparison with non-immersion pupils

Eeva-Liisa Nyqvist, University of Turku

There are two primary goals for this study – first, to analyse definiteness and article use in spontaneous writing in Swedish by 15-year-old Finnish immersion students (n=162) and secondly, to compare their performance with that of non-immersion students at the same age (n=67). Analyses at the group level show that immersion students usually perform significantly better than the control group, but they also reveal similar problems to what L2-Swedish non-immersion students have demonstrated in previous studies, such as omission of indefinite articles and difficulty in choosing the right definite form of the noun. Still, these inaccuracies occurred less often in the data from the immersion students. The studied constructions also show at the group level an acquisition order similar to that reported in previous studies, explainable by different aspects of complexity and cross-linguistic influence. Analyses on the individual level, however, show different acquisition orders depending on the criteria being used.

Finnish and Swedish abstracts at end

Keywords: Immersion, noun phrases, definiteness, article use, Swedish, acquisition order, context-related difficulty, feature-related difficulty

1. Introduction and background

Around 1% of Finnish-speaking pupils in basic education attend early total immersion programmes to learn Swedish as a second language (L2) in Finland (Bergroth, 2015). Studies have shown that immersion students do not acquire native-like L2 proficiency, but they do acquire higher levels of L2 proficiency than non-immersion students (Lyster, 2007). In Finland, up to 81% of immersion students attain higher results in Swedish on their matriculation examinations than the majority of non-immersion students, and they also achieve good results when learning other foreign languages (Bergroth, 2015).

However, immersion pedagogy would benefit from further development (Lyster, 2007). Canadian studies (Genesee, 1987, for review; Harley, 1993, 1998) have pointed to problems with grammatical accuracy (e.g., grammatical gender and verb morphology) among French immersion students. In Finland, immersion research interest has been multifarious (for overview, see Bergroth & Björklund, 2013), but the development of grammatical competence has, so far, gained less empirical attention.

Sipola (2008) reported that grammatical accuracy is the biggest problem for 12-year-old Finnish immersion pupils, who struggle with, for example, noun phrases (NPs) and word order; however, in-depth studies using large corpora to study these topics have not yet been conducted in Finland. Finnish immersion teachers further report that they lack the necessary knowledge about what to concentrate on in Swedish grammar and how to teach it in an optimal way. Thus, it is vital to identify the issues that are problematic for immersion pupils by analysing the production of language from a large group of informants to establish a comprehensive and detailed picture.

The primary purpose of this article is to study how Finnish-speaking, 15-year-old immersion pupils master definiteness and article use in Grade 9 (i.e., at the end of the immersion programme) and what kinds of difficulties they still have with these constructions. According to

Hyltenstam (1992), definiteness and article use (see 1.2 below) cause problems for even advanced learners of L2-Swedish, such as immigrants who have lived in Sweden for decades. Their inaccuracies are rare, but similar to the ones made by less advanced learners. Hence, it is interesting to study whether this category is also problematic for immersion pupils, a very specific group of advanced learners. A multitude of studies in Sweden and Finland have addressed the acquisition of grammar in L2-Swedish. This focus is also relevant in order to study immersion pupils separately from other L2-learners of Swedish, as this especially intensive and sustained educational programme combines both communication and a focus on form in a way that makes it different from many other programme models. The aim of this article, therefore, is not to recommend specific didactic interventions, but rather to identify the problematic structures, so that new didactic interventions can be planned and implemented.

A taxonomy by Housen and Simoens (2016) distinguished among *feature-related* (caused by properties of a linguistic construction), *context-related* (caused by differences in learning conditions), and *learner-related* (individual characteristics) difficulties. My article concentrates on context-related difficulties by comparing a group of immersion pupils to a group of non-immersion pupils. An analysis of the problems encountered by these immersion pupils will offer valuable information about which aspects of definiteness and article use are the most difficult and hence must become the precise focus of more explicit instruction, i.e., on that feature-related difficulty.

2. Background to the study

2.1 Definiteness and article use in Swedish and Finnish

The distinction between definite and indefinite referents is vital in every language, although all languages do not encode this difference morphologically. A referent is definite when *familiar* to the hearer/reader due to the speech situation (e.g., *Can you close the door?*), general knowledge

(e.g. ***the*** *Parliament*), or because it has already been mentioned more or less directly (*e.g., I met a man with two dogs.* ***The*** *dogs were...; I was at a concert.* ***The*** *orchestra was playing Mozart.).* A definite referent can also be *identifiable,* i.e., the listener/reader is capable of matching it with a real-world entity (e.g. ***the*** *President of France*), or it is *unique, i.e.,* there is only one entity that satisfies that offered description (e.g***.*** ***the*** *Moon*, ***the*** *railway station*) (Lyons, 1999).

Count nouns have four forms in Swedish, i.e., indefinite singular and plural (*en bil, bil-ar* "a car", "cars") and the definite singular and plural (*bil-en, bil-ar-na* "the car", "the car-s"). A fifth form is the base form, an indefinite singular without an indefinite article, e.g., *bil* in *Kan du köra bil?* ("Can you drive a car?"), which appears in those contexts where the class of the referent is more important than its individual entity is (Teleman, Hellberg, & Andersson 1999a, 1999b). Indefinite singulars are preceded by an indefinite article, whereas the definite forms have a definite suffix. Definite NPs that are preceded by an adjective attribute (commonly called *double definiteness*) also have a separate definite front article, e.g., *den stor-a bil-en, de stor-a bil-ar-na* ("the big car", "the big cars"), and their adjective attributes also have a definiteness suffix (*-a*), which the indefinite NPs do not have (Teleman et al., 1999a).

The choice of the definiteness form often depends on lexical factors. There is a multitude of attributes in Swedish that will determine the form of the following noun (and also that of the possible adjective attribute) (Teleman et al., 1999a). In the following discussion, I use the umbrella term *PR/G-attribute* when referring to these pronominal/genitive attributes. Several frequent and semantically definite PR/G-attributes (e.g., all possessive and genitive attributes, see examples below) are actually constructed using indefinite nouns, i.e., the relationship between form and meaning thus becomes inconsistent:

Min/Johns glad-a hund-Ø	*denna glad-a hund-Ø*
My/John's happy-DEF dog-INDEF	this happy-DEF dog-INDEF
"my/John's happy dog"	"this happy dog"

Some definite attributes, however, are constructed using definite nouns:

den här glad-a hund-en

this happy-DEF dog-DEF

"this happy dog"

Indefinite pronominal attributes are always constructed using indefinite nouns (Teleman et al., 1999a), i.e., the relationship between form and meaning then becomes clear:

varje glad hund

every happy-INDEF dog-INDEF

"every happy dog"

In Finnish, definiteness is usually interpretable from context, common knowledge, or word order. Like many other languages, Finnish tends to place new information at the end of a sentence and places older information at the beginning (Hakulinen et al., 2004):

Kadu-lla on auto.	*Auto on kadu-lla.*
street-DEF-on is car-INDEF	car-DEF is on street-DEF
"There is a car on the street."	*"The car is on the street."*

Finnish does not contain any linguistic elements that express definiteness only (Vilkuna, 1992); hence L2 learners of Swedish with Finnish as their L1 perceive definiteness marking as redundant (cf. Ringbom, 1987; Sajavaara, 1983). Morphological (case suffixes) and lexical elements, i.e., definite pronouns, *tämä*/*nämä* ("this/these"), *tuo/nuo* ("that/those"), and *se*/*ne* ("it/they"), and the indefinite pronouns *eräs*, *yks(i)* ("a/an," "one"), *joku*/*jotkut* ("some"), and *sellainen* ("such") are used to express definiteness in cases of potential ambiguity, particularly in

spoken language (Hakulinen et al., 2004). Especially *se, tämä,* and *yksi* are often used in a way that resembles article use in the Germanic languages, while native speakers never regard any omission of them as being grammatically inaccurate (Laury, 1996).

2.2 Previous research

2.2.1 *Acquisition of definiteness in L2 Swedish*

The central result from previous research on the acquisition of definiteness in L2 Swedish is that generally all L2 learners of Swedish struggle with the formal aspects of definiteness morphology. Those whose L1 lacks morphological definiteness will also have difficulties when making the choice between the definite and the indefinite form.

According to Axelsson (1994), the acquisition of definiteness morphology advances in three stages. The NPs without definiteness morphology (e.g., the base form, indefinite plural) are acquired first. The second stage with somewhat lower accuracy scores includes NPs with definiteness morphology (definite and indefinite singulars). The role of the L1 is also visible, as Finnish-speaking informants struggled with the indefinite article while they are mastering the definite singular due to the rich inflection system of Finnish NPs. The third level, NPs with adjective attributes, i.e., the most complex forms, yield low accuracy even for advanced learners.

My doctoral thesis (Nyqvist 2013) is so far the most extensive longitudinal study on the acquisition of definiteness and article use by Finnish-speaking L2 learners of Swedish (pupils in comprehensive school from Grades 7–9). These results showed that definiteness in Swedish is a manifold phenomenon: NPs without definiteness morphology are already mastered by the end of the Grade 7, whereas complex NPs still pose problems even in Grade 9. Students have the most problems with the formal aspects of Swedish NPs, while the choice between indefinite and definite meaning is less problematic because they can utilize their previous knowledge of English.

A study on definiteness and article use by adult learners (Nyqvist, 2015) showed, conversely, that the choice of definiteness was a bigger problem than the formal aspects, but that acquisition order on the group level was similar to the order presented by teenagers (Nyqvist, 2013), namely, complex NP types are preceded by the formally simple ones. On the individual level, acquisition order depends on the criterion used, because with systematic use as the criterion, the most frequent NP types seem to be acquired first and the less frequent last, although the less frequent types are not necessarily more difficult.

Definiteness marking has not been a relevant concern for Canadian studies on immersion, as both English and French express definiteness morphologically (World Atlas of Language Structures, 2017). Harley (1993), however, stated that constructions unfamiliar to the learner's L1, along with irregular, low-frequent and non-salient constructions, are difficult for immersion pupils to acquire without more explicit instruction. It appears that these findings are also applicable to definiteness marking for Finnish-speaking learners of L2 Swedish, as Finnish lacks morphological definiteness, and definiteness morphology is difficult to perceive in the input; certain NP types (e.g., double definiteness) are also low-frequency.

2.2.2 *Usage-based grammar and feature-related difficulty*

Because second language acquisition (SLA) during immersion takes place via communication, it can be viewed as a process used for determining linguistic constructions from usage. The constructions are form-meaning mappings that can just as well be words and whole utterances. They undergo different stages of analysis and become categorised as varying grades of abstractness and become gradually rooted as language knowledge (Ellis & Robinson, 2008; Bybee, 2008.)

According to Usage-Based Grammar, language usage affects the cognitive representations of a language. Thus, grammar is an implicit, cognitive organization of a learner's actual language experience, and it develops further, as new constructions are added to the

inventory (Bybee, 2008; Lieven & Tomasello, 2008). This kind of knowledge is procedural, i.e., available for unconscious use (Ellis, 2008). It is acquired through practice, and it is difficult to change afterwards, if practise and repetition lead to inaccuracies (Bybee, 2008; Ellis, 2008). In traditional instruction, the problem is usually the opposite. The learners might have a lot of declarative knowledge, i.e., factual information that they cannot yet use in actual communication in the new language (Ellis, 2008).

SLA in immersion also differs from L1 acquisition: immersion pupils already have acquired a language, so the aim of immersion is to complete their linguistic competence, not replace their L1. Immersion pupils start learning their L2 at an early age and in a similar way as L1 learners do, namely, by hearing the language around them and starting to communicate when they are ready to do so. Hence, it is relevant to draw some parallels between the acquisition of L1 and SLA during immersion learning. Children initially acquire their L1 by repeating concrete utterances that they hear in their everyday lives. Gradually, these form-meaning mappings (constructions) become increasingly complex and schematised. The learners find out more or less consciously the ways in which the parts of the construction link together and thus contribute to their meaning and build relationships between both (Lieven &Tomasello, 2008). Over time, the learners discover regularities in these pairings and start varying them, eventually discovering the abstract formulae behind the concrete utterances. One can, for example, learn Swedish nouns in their definite form and later notice that they all have a certain suffix.

It has, nevertheless, become clear that many aspects of L2 cannot be learned simply by exposure to that L2 (Long, 1990), thereby implying that certain grammatical constructions are more difficult to acquire than others are. *Frequency* plays a crucial part for SLA. Traits that occur more frequently in the input are easier to learn, as that repetition strengthens memory representations and makes them more accessible in one's memory. Extremely frequent sequences can become autonomous, i.e., acquired as wholes, as if they are now independent from a general

pattern, and consequently, these can help the learner analyse similar, less frequent forms (Bybee, 2008).

A disadvantage of frequency is that frequent linguistic elements, such as articles, tend to reduce phonologically, which makes them difficult to notice in new input (Bybee, 2008; Ellis, 2016) and thus difficult to acquire. *Salience*, or how easy it is to notice a construction in any input, is the most important factor (Goldschneider & DeKeyser, 2005), as one cannot acquire what one has not yet noticed (Schmidt, 1990).

Other factors discussed in the literature are *structural complexity*, and this factor involves the choice of several morphemes/allomorphs. This factor is often manifested by *redundancy*, i.e., the same meaning being expressed by several more elements than are semantically necessary. For example, double definiteness is redundant in Swedish, as definite meaning is expressed by one article and two suffixes. Complexity can also occur in the relationship between form and meaning (DeKeyser, 2005); many Swedish NPs with PR/G-attributes have a definite meaning, but they are constructed using indefinite nouns, that is, they are formally rather simple, but L2 learners will find them illogical (Järvinen 2010). An inconsistent relationship is usually problematic for these learners even when the actual construction is frequent. A low-frequent construction, on the other hand, can be learned easily, if the relationship is clear. If the relationship is neither clear nor inconsistent, then its frequency may be the criterion that determines whether acquisition occurs or does not (DeKeyser, 2005).

Transfer from L1 can hinder acquisition if the L1 is less complex than the L2 – as is the case in the definiteness of L2 Swedish for most Finnish-speaking learners (Filipović & Hawkins, 2013). When the meaning of a grammatical trait is novel, abstract, or both, a learner who has problems with that particular meaning also sees the linguistic elements that express it as redundant (DeKeyser, 2005; Jarvis & Odlin, 2000; Jarvis, 2002; Odlin, 2003). This redundancy is especially evident when the learner is obliged to prioritise meaning (e.g., Trenkic, 2007).

As the L1 makes certain grammatical elements redundant for the learner, it can also be assumed that the L1 of the learner influences what the learner finds salient, thereby also explaining why the definite singular is easier than the indefinite one for Finnish-speaking L2 learners of Swedish. The suffixes are more salient than the articles, as they do occur in the L1 of the learner (Axelsson, 1994; Nyqvist, 2013, 2015). The hindering effect of the L1 decreases, however, when acquisition advances (e.g., Bybee, 2008; Jarvis, 2002; Trenkic & Pongpairoj, 2013). At the same time, it is important to remember that difficulties may simultaneously have several sources: L1 certainly plays a part in the fact that Finns repeatedly omit articles, but the fact that these articles are lacking in salience also plays a role in that omission.

Previous studies (e.g., Ringbom, 2007) have shown that learners are likely to transfer elements from another L2 when there is a substantial typological distance between the L2 and the L1, but a lesser distance between the L2 and the target language. Because English is the most common first L2 learned by non-immersion pupils (Finnish National Board of Education, 2014), it is likely to play a central role in their acquisition of the L3 Swedish. Non-immersion pupils had fewer problems with the choice of definiteness than they had with the formal aspects of definiteness marking (Nyqvist, 2013), and that implies that they were able to take advantage of their knowledge of English. The surface structure of Swedish and English NPs are, however, so different that knowledge of English cannot help them with formal aspects of the Swedish NPs (Nyqvist 2013, see also Ringbom, 2007). Conversely, immersion pupils learn Swedish as a L2, but even they live in continuous contact with English; hence, their L2 Swedish might bear certain traits of their L3 English.

In the following, I describe the informants, the data and methods used in this study and analyse the data on both group and individual level.

3. The Study

3.1 Study informants

Informants were Finnish 9th graders (n=162) enrolled in Swedish immersion (henceforth IM) and Finnish 9th graders enrolled in non-immersion, traditional instruction (henceforth TR) (n=67). Until 2016,[1] the majority of Finnish-speaking non-immersion pupils started learning Swedish at the age of 13 with two weekly lessons (Finnish National Board of Education, 2014; see also Finnish National Agency for Education, 2014). Even though Swedish is one of the official languages of Finland, the pupils in non-immersion settings learn Swedish, *de facto*, as a foreign language: course materials and teachers are their principal sources of input as they usually lack everyday contact with Swedish. The Finnish immersion pupils, conversely, learn Swedish mainly as a result of communication; however, 55 % of my informants reported that they *do not* use Swedish outside school, in their activities or their neighbourhoods.

Both informant groups started learning English at the age of nine. Hence, both groups received as much instruction in English in terms of time; however, the immersion pupils learned Swedish as their L2, whereas the TR students learned Swedish as a third language (L3). My doctoral thesis (Nyqvist, 2013) demonstrated that learners of L3 Swedish can take advantage of their knowledge of English when expressing definiteness. The Swedish system for marking definiteness, as explained above, differs profoundly from that of English, so English rules for definiteness cannot help when learning Swedish and vice versa.

3.2 Data collection

The primary data from IM students consist of spontaneous writing (circa 200-word narratives, entitled *My Dream Journey/ Holiday*), and shorter e-mail tasks where the informants asked for

[1] Beginning autumn term 2016, the obligatory instruction in Swedish for Finnish-speaking pupils now starts in 6th grade at the latest.

information about homework and school activities while away from school. The control data from TR students consist of 60–70 word narratives on the informant's plans for the future (informants for my doctoral thesis, Nyqvist, 2013).

3.3 Research questions and hypotheses

As there are no previous large, in-depth studies on Swedish immersion grammar learning, these current hypotheses on the acquisition of definiteness are based on results acquired from non-immersion Swedish learning in Sweden and in Finland (see 2.2). The specific research questions and hypotheses thus are the following:

- RQ1: Are there any differences in frequencies of the different types of NPs between IM and TR?
    - H1: The most common NP types have rather similar frequencies seen in both sets of data, but the complex NPs (e.g., double definiteness) have higher frequencies in IM (Nyqvist, 2013, 2015).
- RQ2: Do the majority of the inaccuracies in IM involve the choice of definiteness (i.e., the indefinite form in constructions that require the definite form, henceforth *definiteness inaccuracies*) or do they involve the form of the NP (i.e., omission of an article, henceforth *formal inaccuracies*)?
    - H2: Formal aspects pose more problems than the choice of definiteness, which has also been the case in most of the previous studies that have used Finnish-speaking informants (Axelsson, 1994: Nyqvist, 2013).
- RQ3: What kind of differences in accuracy appear between the IM and TR groups?
    - H3: Accuracy scores are higher in IM, but the inaccuracies are qualitatively similar in both groups, as they have been in previous studies (Hyltenstam, 1992).
- RQ4: What kind of acquisition order is there in IM?

- H4: The acquisition order on the group level is largely similar to that seen in previous studies, i.e., the most complex NPs are acquired last (Axelsson, 1994: Nyqvist, 2013). On the individual level, the order based on a percentage criterion resembles those seen on the group level, whereas the order based on systematic use favours frequent NPs that appear to be acquired first (Nyqvist, 2015).

3.4 Data analysis

Data analysis was founded on obligatory occasions (Ellis & Barkhuizen, 2005), i.e., constructions that the informants should use in line with *Svenska Akademiens Grammatik* ("The Swedish Academy Grammar", Teleman et al., 1999a, 1999b). The writing tasks were planned so that informants could produce concrete language wherein the obligatory occasions would thus be easy to identify. On the group level, data were analysed from the point of view of frequencies as well as accuracy. Frequencies were calculated by dividing the number of a certain NP type by the number of all NPs, and the accuracy scores were obtained by dividing the number of accurate obligatory occasions of an NP type by the total number of obligatory occasions of that same NP type. The statistical software package, SPSS, was used to calculate the statistical significance[2] of the differences in accuracy scores between the two groups. Acquisition sequences were established in line with Brown's principle (1973), wherein an acquisition sequence delivers a difficulty hierarchy: high accuracy implies early acquisition and consequently points to an easy construction (see also Collins, Trofimovich, White, Cardoso & Horst, 2009).

On the individual level, the analysis concentrated on the most fundamental NP types and established the acquisition sequences using implicational scales that can provide information about how the acquisition of the NPs developed in the learning of individual informants. The

[2] Pearson's chi-square, value of $p > 0.05$

method shows whether the acquisition of, for example, the indefinite form singular, actually implies a mastery of, for example, the definite form singular. In that case, it can be concluded that the definite singular is acquired before the indefinite singular is acquired. As these scales often show, some of the informants will master a difficult form before they master an easier one (Ellis & Barkhuizen, 2005). Herein the statistical validity of the scales was established by calculating the *coefficient for reproducibility* and *the coefficient of scalability* (the abbreviations C_{rep} and C_{scal} in Tables 2 and 3 below, the limiting values being 0.9 and 0.6, respectively; Hatch & Lazaraton, 1991).

The acquisition of the studied constructions was operationalized as implicational scales in two ways. The criterion for *systematic use* measures emergence instead of mastery and does vary in different studies, as each researcher has to decide how many occurrences are required to establish the pattern of a form emerging in the production. That criterion is often defined as having at least three accurate occurrences of a construction per text (Pallotti, 2007). It is also the criterion used in this study.

Missing data, i.e., cases where a learner used a construction only 1–2 times accurately are handled differently in different studies. Rahkonen and Håkansson (2008) did not automatically regard missing data as a sign of non-mastery, and they instead chose to disregard missing data when calculating the validity of their scales. The coefficient of reproducibility was thus calculated herein using a modified version of Guttmann's formula, namely, missing data are subtracted from the product of the *number of subjects* times the *number of items* before the *number of errors* is divided by the result of that math operation (cf., Hatch & Lazaraton, 1991). As systematic use is about emergence (and not mastery) and IM represents advanced learners, it is unlikely that any missing data may be due to the fact that a certain form has not emerged in their L2. Hence, Rahkonen and Håkansson's (2008) version of the formula as noted below was utilised for the current study:

$$C_{rep} = 1 - \frac{\text{number of errors}}{((\text{number of subjects})(\text{number of items}) - \text{missing})}$$

The *percentage criterion* involves choosing a percentage level of accuracy, but even this choice is arbitrary (Pallotti, 2007). Here I set the percentage criterion at 80%, as this criterion is often used in linguistic studies (e.g., Rahkonen & Håkansson, 2008; Glahn et al., 2001). The strength of this criterion lies in the fact that it does not favour frequent constructions like systematic use does, and missing data also do not occur. The fact that only one accurate occurrence is enough to fill the criterion may, on the other hand, lead to situations where the criterion can be met by simply accident.

4. Results – Group Level Analysis

This section presents the central results of the study on a group level by displaying the frequencies and accuracy scores for the 6,445 NPs produced by IM students (853 by TR pupils). Previous research (e.g., Axelsson, 1994; Nyqvist, 2013) has shown that an adjective attribute often affects the accuracy score considerably, so it is relevant to distinguish between NPs with and without adjectives. The normative analysis presents the most typical inaccuracies. The following discussion illustrates that analysis, using authentic examples gleaned from the data. The actual inaccuracies are marked with an asterisk (*). The definiteness suffixes have for the sake of clarity been separated from the word stem by a hyphen (-), and the correct form is given in parentheses after the inaccuracy examples.

4.1 *Frequency analysis*

Figure 1 includes all NPs with a frequency ≥ 1%[3] in IM and the equivalent frequencies in TR. IM also includes five inaccurate NPs (2 in TR) that could not be analysed.

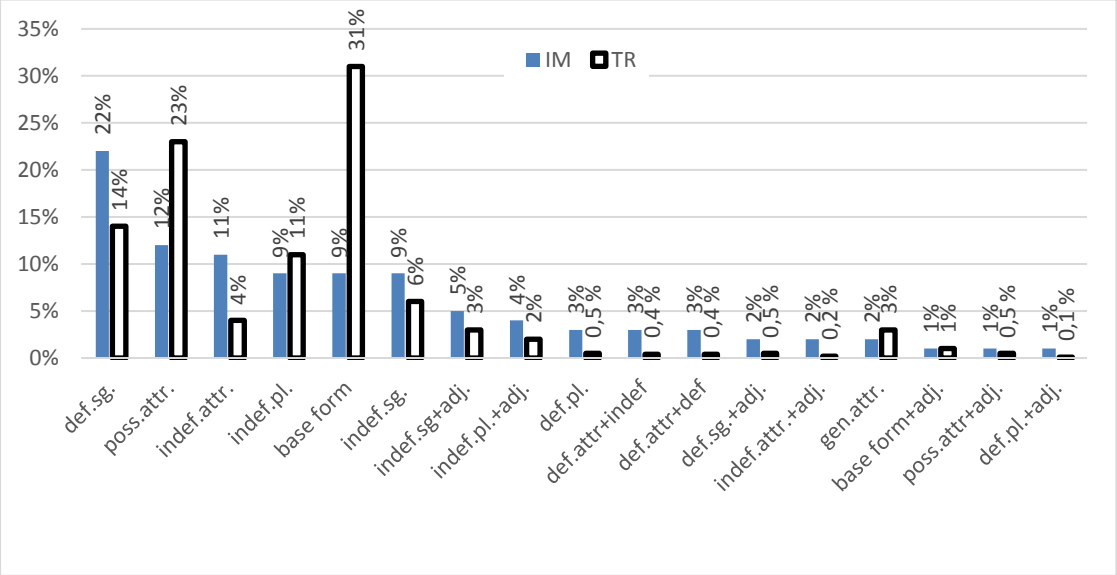

Figure 1: Frequencies of the most common NP types in IM and TR

As Figure 1 shows, certain NP types dominate in IM, and the majority have low frequencies. Definite singulars cover more than 20% of all NPs, while NPs with possessive and infinitive attributes, base forms, and indefinite plurals and singulars (all without an adjective attribute) have a frequency of around 10% each. Although different themes in IM and TR[4] do lead to partially different needs for expression, the frequencies are rather similar in both sets of data, even though more complex NPs are more frequent in IM. Hence, H1 holds here.

The dominance of definite singulars in IM is comprehensible, as they are mostly used to refer to concepts that belong to the lives of the informants (e.g., *sommarlov-et*, "the summer

[3] NPs with a frequency <1% are PR/G-attributes with adjectives, and the abbreviations *def.attr+indef* and *def.attr+def* stand for definite attributes that are constructed using indefinite/definite attributes, respectively.
[4] For a detailed analysis of TR, see Nyqvist 2016.

holiday") and are relevant for the trip/holiday they are describing (e.g., *ett stort simbassäng i bakgård-en*, "a big swimming pool in the backyard"), or they refer to the identifiable (e.g., *strand-en,* "the beach"*; hav-et* "the sea"), or unique referents (e.g., *sol-en*, "the sun").

Possessive attributes are frequent, as the IM informants wrote about their dreams and plans concerning a holiday (e.g., *min vänner* "my friends", *min familj* "my family"). However, they are less frequent than in TR, as the informants are better able to vary their production and are not "locked" into a single manner of expression, as the informants in TR might be. Indefinite attributes are rather frequent in IM (10%), as one of their subcategories, interrogative attributes, is common in e-mail tasks, e.g., *Vilka uppgifter har jag att göra?* ("Which exercises should I write?").

The holiday theme of IM prompts utterances with typical contexts as a base form, e.g., months: *i maj* ("in May"), food and drink: *dricka kaffe* ("drink coffee"), and NPs where the class of the referent is more important than the individual entity: *konst, mode* ("art", "fashion"). However, due to the themes of these texts, contexts are substantially more frequent in TR writing.

4.2 *Normative analysis*

The following section first presents the accuracy scores among IM students and then compares them to each other and to the TR students'. Secondly, it describes the most characteristic problems seen for indefinite and definite NPs. The data include 615 inaccurate NPs; less than 10% of the NPs were inaccurate (15% in TR, Nyqvist, 2013). Thus, there were significantly fewer inaccuracies in IM than in TR ($p=0.000$). The inaccurate NPs in IM totalled 665 of which 60% concerned form (56% in TR, Nyqvist, 2013). Formal inaccuracies were again significantly more frequent than definiteness inaccuracies ($p=0.000$), and thus H2 holds true here. Figure 2

shows the accuracy scores for the different NP types with a frequency >1%. As adjectives lack definiteness morphology in indefinite forms, their forms are considered only for definite NPs.

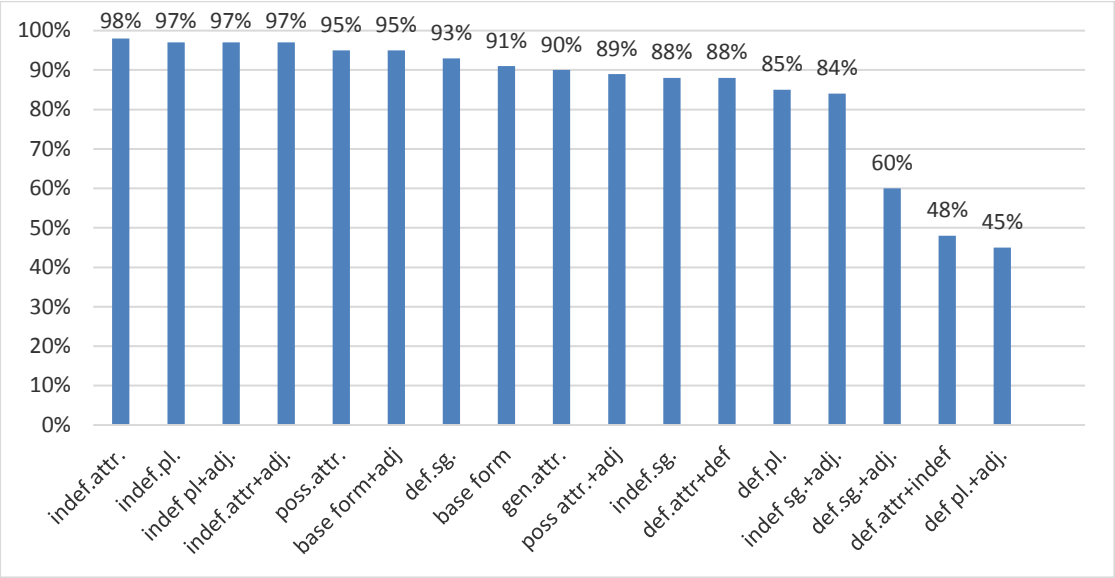

Figure 2: Accuracy scores for the most common NP types in IM

As Figure 2 shows, IM accuracy scores exceeded 80% for most NP types. The ones with the lowest scores (≤60%) were the most complex, i.e., double definiteness in the singular and plural (formally complex) and definite attributes constructed using indefinite nouns (a complex relationship between form and meaning) (DeKeyser 2005). These also demonstrated a low-frequency in the data and probably also in the input the informants received.

A central trait in TR is the clear differences seen in accuracy for NPs with no definiteness marking, for NPs with definiteness suffix, and for NPs with an article or two suffixes, i.e., the least complex NPs are acquired before the most complex ones (cf., Nyqvist, 2013). In IM, the accuracy scores do not show similar and visually clear "stages", but still, a lowering accuracy hierarchy equals a rising complexity hierarchy. Hence H4 holds, for this part, true. The indefinite plurals are again significantly more accurate than the indefinite singulars and the definite singulars and plurals (p=0.000 in all cases). Definite singulars, in turn, are significantly easier

than definite plurals (p=0.000), i.e., one suffix is easier than two. Definite and indefinite singulars are equally complex, as they include one grammatical morpheme, while the definite singular is significantly easier (p=0.001). This phenomenon occurred also in TR and in Axelsson (1994), and it is explainable by the fact that Finnish has suffixes, but not articles, and hence the form with a suffix is easier for Finnish L2 learners of Swedish (Axelsson, 1994; Nyqvist, 2013, 2015). The definite forms with an adjective are among the most difficult, so these accuracies are significantly lower than the accuracies for equivalent forms without an adjective, (p=0.000). In short, acquisition order is similar in both IM and TR.

For the PR/G-attributes, the possessive attributes are used significantly more accurately than the genitive attributes and the definite attributes that are constructed using indefinite nouns (p=0.017 and p=0.000), although all NP types are equally complex as far as the relationship between form and meaning is concerned. Possessive and genitive attributes also resemble each other semantically, and they are probably equally frequent in the input. Accuracy differences are explainable by the fact that possessive attributes comprise a limited number of grammatical words, usually learned at an early stage, whereas every genitive attribute must be built productively, by adding an *-s* to an NP. The semantic similarity of these genitive attributes to the possessive attributes may also explain why genitive attributes are significantly easier (p=0.000) than definite attributes constructed using indefinite nouns; the last-mentioned attributes indeed build a miscellaneous group of pronouns that must be learned by heart.

Table 1 summarizes the most central accuracy differences between IM and TR students. When the accuracies in IM were compared to those found in TR, scores for the immersion pupils were in most cases significantly higher.[5] Hence, one can assume that learning in the immersion

[5] However, if a Bonferroni correction is made, indefinite plurals and indefinite singulars with an adjective attribute are no longer significant, as the p-value is >0.05/8 (where 8 is the number of the conducted tests).

program favours the acquisition of definiteness and article use; thus H3 holds at least partially, as the accuracy scores do tend to be higher among IM students.

Table 1: Accuracy scores for the most frequent NP types in IM and TR

NP	*accuracy in IM*	*accuracy in TR*	*p*
Indef. pl.	97%	93%	0.044
Base form	91%	91%	0.92
Poss. attr.	95%	85%	0.000
Def.sg.	93%	84%	0.000
Indef. pl+adj.	97%	76%	0.000
Indef. sg.	88%	66%	0.000
Indef. sg + adj.	84%	64%	0.017
Def.pl.	85%	60%	0.14

4.2.1 *Indefinite NPs*

In *base forms*, a majority of the inaccuracies concerned the choice of definiteness, as the definite singular is used in context for the base form. In *Jag älskar *teater-n* (*teater;* "I love theatre"), the informant may overuse the definite form, because the rich inflectional system of her L1 makes the definiteness suffix accessible to her (cf., Axelsson, 1994). Another possible explanation is that she assumes that *teater* belongs to general knowledge, and therefore, it has to have a definite form. This is a rather low-frequent type of definiteness inaccuracy (18% of all the definiteness inaccuracies), but native speakers do find it disturbing, as the definite form leads to an unnecessary memory search by the listener/reader; the definite form implies that s/he should know which referent one is talking about (Nyqvist, 2014).

Definiteness inaccuracies also dominated in *indefinite plurals* in both TR (Nyqvist, 2013) and IM, where up to 91% of the inaccuracies consisted of overuse of the definite form in contexts where using the indefinite form was obligatory, e.g., *Jag skulle vilja se *sevärdheter-na* (*sevärdheter;* "I would like to see tourist attractions"). For the *indefinite singulars*, 84% of the inaccuracies were formal. Omission of the indefinite article, e.g., *Vi skulle åka till *paradisholm*.

(*en paradisholm*; "We would go to a paradise island") was the most common problem, as it was in TR. This is a notorious problem for Finnish-speaking learners of Swedish, as the indefinite article does not occur in Finnish, and hence, it is viewed as redundant by Finns (Jarvis & Odlin, 2000). The article is additionally not salient in the input and has an abstract meaning (DeKeyser 2005). Omission of the indefinite article is, however, significantly less common in IM than it is in TR ($p = 0.000$ in NPs without adjectives, $p = 0.017$ in NPs with adjectives), which suggests that influence of the L1 was decreasing for the immersion pupils (cf., Jarvis, 2002; Trenkic & Pongpairoj, 2013).

4.2.2 Definite NPs

In a similar way as in TR, the majority (97%) of inaccuracies *in definite forms* were definiteness inaccuracies. In *singulars*, the base form was overused instead of the definite one: *Min dröm är att åka till Miami (…) Sen går vi till *flygplats och till *flygplan.* (*flygplats-en, flygplan-et;* "My dream is to go to Miami (…) Then we go to the airport and wait for the plane").

In *plurals*, the indefinite plural was overused, e.g., *Kan jag använda mina egna åsikter i *övningar?* (*övningar-na*; "Can I express my own viewpoints in the exercises?"). Altogether, these covered 67% of all definiteness inaccuracies and occurred in most cases when the definite form lacked a direct antecedent, i.e., the same noun was not mentioned earlier.

In TR, overuse of the base form/indefinite form was explicable due to the fact that the omission of grammatical morphemes is common in beginner texts, and Finnish materials for L2 Swedish overemphasize NPs with an antecedent, thus applying the rule, "Use the definite form when the noun has been mentioned earlier" (Nyqvist, 2013), although definite NPs in Swedish in most cases lack an antecedent (Fraurud, 1988). Nevertheless, it is highly unlikely that immersion pupils would have been exposed to such repetitive input. Hence, it is possible that overuse of the

base/indefinite forms is caused by the fact that Finnish-speaking learners fail to recognize this kind of context in their definite forms, as their L1 lacks that morphological definiteness.

Double definiteness had a rather low accuracy score in both the singular and the plural, and it has been shown in previous studies (notably Axelsson, 1994) to be problematic for even advanced learners of L2 Swedish. Immersion pupils have similar problems with this form just as the TR pupils did. These problems are explainable by two factors: Double definiteness is a very complex NP type, and also has low frequency in the input. Consequently, not even immersion pupils experience enough occasions to practise these NPs.

NPs with double definiteness often include several inaccuracies. In the singular, the most common problem is the omission of the definite front article, e.g., **blå-a hav-et* (*det blå hav-et*, "the blue sea"); however, overuse of the base form, e.g., **andra *land* ("the second country") is also common. In the plural, in contrast, overuse of the indefinite form, e.g., *de stora *byggnader* (*de stora byggander-na*, "the big buildings"), is more frequent than is omission of the definite front article, e.g., **bästa saker …* (d*e bästa saker-na*, "the best things"). Here again, around 50% of NPs with double definiteness and overuse of bare/indefinite forms in both the singular and the plural are antecedent-less.

As the PR/G-attributes determine the choice of the definiteness form of the noun, all inaccuracies in these NPs concerned the formal aspects. The archetypal inaccuracy is overuse of the definite form, i.e., the informants tend to be "logically consistent" and match the semantic relationship (cf., DeKeyser, 2005) by choosing the definite form, e.g., *mitt *favoritishockeylag-et* (*mitt favoritishockeylag*, "my favourite ice hockey team"), *Amerikas östra *kust-en* (*Amerikas östra kust* "America's East Coast"), *denna fina stora *stad-en* (*denna fina stora stad*, "this great town"). This inaccuracy is most common with the definite attributes, which are less frequent than the possessive and genitive attributes. In my doctoral thesis, this inaccuracy was significantly more frequent in Grade 9 than in the earlier grades, which indicated that the

informants had become increasingly more aware of the definite meaning of NPs and simply jumped to the conclusion that definite meaning automatically entails using a definite form (Nyqvist, 2013).

5. Results – Individual Level Analysis

This section presents analyses of the use of *base forms, indefinite, and definite forms in the singular and in the plural,* and *singular NPs with possessive pronouns* on an individual level. The frequencies of most of these were ≥ 10% at the group level when the adjective attributes were disregarded. As more than 70% of NPs with possessive attributes have singular nouns, only singular NPs were taken into account here.

Systematic and accurate use at the 80% criterion is marked in Tables 2 and 3 with a plus (+), whereas missing data (in systematic use) is marked with a slash (/). The slashes have been placed inside the vertical line when a more difficult form has been mastered (as in the 2nd row of Table 2: missing data in the base form, but systematic use in the definite plural). The NPs that were lacking systematic use or accurate use at the 80% criterion are marked with a minus (-). The exceptions from the ideal implicational scales, i.e., a more difficult NP is mastered despite the fact that an easier one is not, are noted by an exclamation mark (!).

In the tables, the studied constructions are horizontally ordered from the easiest to the most difficult (in other words, the NP with most informants using it systematically/at an 80% criterion stands farthest to the left), and the learners are vertically ordered according to the number of constructions for which they met the criterion. Due to the large number of informants, the rows in the tables do not represent individual learners, but rather groups (abbreviated *gr*. in the tables below) of learners – all informants using a certain type of NP systematically/at the 80% criterion according to the same pattern are shown in the same row.

5.1 Systematic use

Table 2 presents evidence of systematic use of the central NP types in IM students. A salient trait of Table 2 is its great variation, as 64% of the rows here represent only one informant.

Table 2: Systematic use of the central NP types in IM

gr.*	**def. sg.**	**indef. pl.**	**indef. sg.**	**poss. sg.**	**base form**	**def. pl.**	**in total**	**%**	**deviations**
1	+	+	+	+	+	+	8	5%	0
2	+	+	+	+	/	+	5	3%	0
3	+	+	+	/	/	+	1	<1%	0
4	+	/	+	/	/	+	3	2%	0
5	+	+	+	/	+	+	2	1%	0
6	+	+	/	+	+	+	4	2%	0
7	+	+	/	/	+	+	2	1%	0
8	+	/	+	+	+	+	2	1%	0
9	+	+	+	+	+	/	22	14%	0
10	+	+	/	+	+	/	1	<1%	0
11	+	/	+	+	+	/	1	<1%	0
12	+	+	+	/	+	/	5	3%	0
13	+	/	+	/	+	/	1	<1%	0
14	+	/	/	+	+	/	1	<1%	0
15	+	+	/	/	+	/	3	2%	0
16	/	/	+	+	+	/	1	<1%	0
17	+	+	/	/	+	/	1	<1%	0
18	+	/	+	+	+	/	2	1%	0
19	/	+	+	+	+	/	1	<1%	0
20	+	+	+	+	+	-	15	9%	0
21	+	+	+	/	+	-	8	5%	0
22	+	+	/	+	+	-	1	<1%	0
23	+	+	/	/	+	-	3	2%	0
24	+	/	+	+	+	-	3	2%	0
25	+	/	/	/	+	-	2	1%	0
26	/	/	+	+	+	-	3	2%	0
27	+	/	+	/	+	-	1	<1%	0
28	/	/	/	/	+	-	1	<1%	0
29	/	+	+	/	+	-	1	<1%	0
30	/	+	+	+	+	-	1	<1%	0
31	+	+	+	+	/	/	10	6%	0
32	+	+	/	+	/	/	1	<1%	0
33	+	/	/	+	/	/	1	<1%	0

34	+	+	+	+	/	-	12	7%	0
35	+	+	/	+	/	-	2	1%	0
36	+	/	/	+	/	-	1	<1%	0
37	/	/	+	+	/	-	1	<1%	0
38	/	+	/	+	-	-	1	<1%	0
39	+	+	+	+	-	/	1	<1%	0
40	+	+	+	/	/	/	3	2%	0
41	+	+	+	/	-	+!	1	<1%	1
42	+	+	+	/	/	-	4	2%	0
43	+	+	+	-	/	+!	1	<1%	1
44	+	+	+	/	-	/	1	<1%	0
45	+	+	+	-	/	-	1	<1%	0
46	+	+	+	-	-	/	1	<1%	0
47	+	+	+	/	-	/	1	<1%	0
48	+	+	+	-	+!	-	1	<1%	1
49	+	/	+	-	-	/	1	<1%	0
50	+	+	/	-	-	-	1	<1%	0
51	+	+	-	+!	+!	/	1	<1%	2
52	+	+	-	+!	/	-	1	<1%	1
53	+	+	-	+!	+!	/	1	<1%	2
54	+	+	-	+!	+!	+!	1	<1%	3
55	+	+	-	+!	+!	/	1	<1%	2
56	+	+	-	+!	+!	-	1	<1%	2
57	+	-	/	+!	+!	/	1	<1%	2
58	+	/	/	-	/	/	1	<1%	0
59	/	-	+!	/	-	-	1	<1%	1
60	-	/	/	+!	+!	-	1	<1%	2
61	-	+!	+!	+!	/	-	1	<1%	3
	159	**132**	**127**	**110**	**104**	**68**	**162**	**100**	**23**

C_{rep} = 0.97; C_{scal} = 0.97

*Abbreviations are as follows: def. sg. = definite singular, indef. pl. = indefinite plural, indef. sg. = indefinite singular, poss. sg. = possessive singular, def. pl. = definite plural

As Table 2 shows, 149 informants (92%) followed an ideal implicational scale. Only eight informants used all forms systematically (gr.1). Irrespective of 23 deviations, there were implication relationships between the forms, meaning that the scale is valid. Hence, one can establish an implicational order, namely, definite singular<indefinite plural<indefinite singular<possessive singular<base form<definite plural. The implicational order differs from the acquisition order at the group level, as definite and indefinite singulars are mastered before less

complex forms are. This order is explainable by the fact that systematic use combines accuracy and frequency, and thus that use favours the high-frequent forms (Nyqvist, 2015). Hence the H4 holds true here.

The definite singular was significantly more often systematically used than all other forms except for the indefinite plural (p=0.001 in the indefinite and possessive singular and p=0.000 in base form and definite singular) because of its high frequency (25% of NPs) and accuracy. The indefinite plural was the second most systematically used form, because of its extremely high accuracy score. Both the definite singulars and the indefinite plurals are also frequent in the classroom language, which explains their early emergence. The base form is significantly less often systematically used than the more complex forms (p=0.000 in the definite singular and indefinite plural and singular, p=0.001 in possessive singular), because it has rather low frequency in the data; up to 45 informants had 0-2 occurrences of the base form. Thus, these results should not be interpreted as if the informants had problems with the use of the base form (accuracy 91% on the group level). The definite plural was, on the contrary, significantly less often used systematically than the other forms (p=0.000 in all cases) because it is both low-frequent and formally complex.

5.2 The percentage criterion

Table 3 builds on an 80% criterion and shows slightly less variation than Table 2 did, as 39% of these rows represent only a single informant.

Table 3: Indicates the 80% criterion for the central NP types by IM students

Gr.	indef. pl.	poss. sg.	base form	def. sg.	indef. sg.	def. pl.	in total	%	deviations
1	+	+	+	+	+	+	38	23 %	0
2	+	+	+	+	+	-	32	20 %	0
3	+	+	+	+	-	+!	9	6 %	9

4	+	+	+	+	-	-	15	9 %	0
5	+	+	+	-	+!	+!	2	1 %	4
6	+	+	+	-	-	+!	2	1 %	2
7	+	+	+	-	+!	-	6	4 %	6
8	+	+	+	-	-	-	6	4 %	0
9	+	+	-	+!	+!	+!	8	5 %	24
10	+	+	-	-	+!	+!	2	1 %	4
11	+	+	-	-	+!	-	2	1 %	2
12	+	+	-	+!	-	+!	1	<1%	2
13	+	+	-	+!	+!	-	8	5 %	16
14	+	+	-	+!	-	-	4	2 %	4
15	+	-	+!	-	+!	-	5	3 %	10
16	+	-	+!	+!	-	-	3	2 %	6
17	+	-	+!	+!	+!	-	2	1 %	6
18	+	-	-	+!	+!	-	1	<1%	2
19	+	-	+!	+!	+!	+!	2	<1%	8
20	+	-	+!	-	-	-	1	<1%	1
21	+	-	+!	-	-	+!	1	<1%	2
22	+	-	-	+!	+!	+!	1	<1%	3
23	+	-	-	+!	+!	-	1	<1%	2
24	+	-	+!	-	+	+	1	<1%	3
25	-	+!	-	+!	+!	-	1	<1%	3
26	-	+!	+!	-	+!	-	1	<1%	3
27	-	+!	+!	+!	+!	+!	1	<1%	5
28	-	+!	+!	-	+!	+!	1	<1%	4
29	-	+!	+!	+!	+!	-	2	1 %	8
30	-	+!	-	+!	+!	-	1	<1%	3
31	-	+!	+!	-	+!	-	1	<1%	3
32	-	+!	-	-	-	-	1	<1%	1
	153	**144**	**131**	**131**	**119**	**68**	**162**	**100**	**146**

C_{rep}=0.85; C_{scal}=0.84

Table 3 shows that 91 informants (56%) followed the ideal scale. Up to 23% fulfilled the 80% criterion in all forms. The deviations were, however, so many that the scale becomes invalid. Nevertheless, one can see that the scale follows a similar acquisition order as at the group level: the criterion is most often fulfilled in forms that are lacking definiteness marking, while the definite singular is mastered at the same level as the base form. Definite plurals again fulfil the criterion significantly less often than the other forms ($p=0.000$ in all cases).

6. Discussion

The primary goal of this paper has been to study how the acquisition of definiteness and article use, both notorious sources of difficulty for Finnish-speaking learners of L2 Swedish, has proceeded in the written narratives of 15-year-old Finnish immersion pupils, and further, what kind of NPs still pose problems for them at the end of their immersion learning. These data were then compared to the data gathered from 15-year-old, non-immersion (TR) learners of Swedish. My hypotheses were as follows:

1. The frequencies are very similar in both sets of data, but the most complex NP types are more frequent in IM than in TR.
2. The majority of the inaccuracies concern formal aspects of the NPs.
3. The accuracy scores are higher in IM but the inaccuracies are qualitatively similar in both sets of data.
4. The acquisition orders are on the group level largely similar in both sets of data, whereas the orders on the individual level depend on the criterion being used.

The frequencies were found to be mostly similar in IM and TR students, while complex NPs were more frequent in IM. This might depend on two different reasons: the narratives by immersion students were, naturally, longer than the ones by the control group and hence there

were more natural contexts for these NPs. It is also possible that the pupils of the control group had actively avoided the most complex NPs they had found very difficult to use.

The majority of inaccuracies in IM were generally formal inaccuracies, such as the form of nouns following PR/G-attributes and omission of articles. Omission of the indefinite article is a classic inaccuracy in the L2 Swedish of learners whose L1 lacks articles. However, previous studies show that these inaccuracies also depend on the fact that indefinite articles are not salient in the input, that is, they are inherently difficult, irrespective of their high frequency. Furthermore, their meaning is not conspicuous for Finnish-speaking informants. They are, however, mastered significantly better in IM than in TR, as the role of the L1 decreases in IM students because SLA has advanced.

In PR/G-attributes, the problems depend on the inconsistent relationship between form and meaning. Attributes with definite meaning are constructed using indefinite forms and that is confusing for learners, although the problem is less prominent with frequent attributes such as the possessive ones.

There were fewer definiteness inaccuracies. A central problem in IM is related to overuse of a base form in the antecedent-less definite NPs. Another problematic issue concerns overuse of the definite form instead of using the base/indefinite form, which native speakers can find disturbing. NPs with double definiteness include problems in both the formal aspects and the choice of the right form of the noun. This NP type is, however, so low in frequency that it needs to be studied separately, using its own elicited data. In summary, the inaccuracies are explainable by the characteristics of Swedish (notably salience and complexity) and to some extent also the L1.

Many central accuracy scores were significantly higher in IM than they were in TR. These results suggest that the immersion approach, given the early age of onset, rich input, and abundance of possibilities for communication, does favour the acquisition of definiteness and

article use. The remaining inaccuracies were qualitatively similar to the ones found in TR, but lower in frequency (Hyltenstam, 1992).

The acquisition order in IM resembles the findings obtained in previous studies and then proceeds on the group level from NPs without any definiteness morphology to NPs with a suffix or an article. The complex, definite NPs are also acquired last by immersion pupils. The factors that affect the accuracy of the studied NPs above all concern complexity (of form and the relationship between form and meaning), although even transfer from the L1 plays a role, as the informants mastered definite singular forms significantly better than they did indefinite singulars; in other words, suffixes are easier for them than indefinite articles. It is also possible that the characteristics of the L1 make definiteness suffixes more salient than indefinite articles for these Finnish-speaking informants.

On the individual level, acquisition orders are different, depending on the criterion used. Systematic use shows whether a form has emerged in the L2, whereas the percentage criterion shows to what extent a form is being mastered. When the same number of occurrences is demanded, but different forms have different numbers of obligatory occasions in spontaneous data, certain forms seem to emerge to a greater extent than others do. Therefore, the frequent definite singular is acquired before the less complex constructions, even though the less frequent forms are not necessarily more difficult.

Using the 80% accuracy as the criterion leads to an order similar to the one at the group level. The scale is invalid due to the high number of deviations. Another disadvantage with using percentage as a criterion is that coincidence may play a role, as only one occurrence is enough to fulfil the criterion. However, a commonality for both criteria is that definite plurals fulfil the criterion significantly more seldom than for the other forms. Definite plurals are more difficult than most of the forms even at the group level, but their difficulty also stands out at the individual level.

Even though the majority of NP types are mastered at a high level by the end of immersion, it is important to consider how to help these pupils over the final obstacles in their acquisition of indefinite articles, nouns after the PR/G-attributes, double definiteness, and antecedent-less definite forms and definite plurals. Teachers who are teaching immersion pupils any school subject need to be made more aware of the problems that their pupils can encounter and try to give them more corrective feedback during the lessons (see Lyster & Ranta, 1997). My suggestion for future research is further didactic study on how best to utilise these current results in practical terms and how to address the remaining difficulties that were discovered in this current study.

References

Axelsson, M. (1994). *Noun phrase development in Swedish as a second language: A study of adult learners' acquiring definiteness and the semantics and morphology of adjectives*. Stockholm: Stockholm University Centre for Research and Bilingualism.

Bergroth, M. (2015). *Kotimaisten kielten kielikylpy. Vaasan yliopiston julkaisuja. Selvityksiä ja raportteja 202* [Immersion of the National Languages]. Vaasa: University of Vaasa.

Bergroth, M., & Björklund, S. (2013). Kielikylpyohjelman tutkimustuloksia Suomessa. [Results from immersion studies in Finland] In L. Tainio & H. Harju-Luukkainen (Eds.), *Kaksikielinen koulu – tulevaisuuden monikielinen Suomi/Tvåspråkig skola – ett flerspråkigt Finland i framtiden* [Bilingual school – a multilingual Finland of the future]. (pp. 91–114). Jyväskylä: Finnish Educational Research Association. Kasvatusalan tutkimuksia 62.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

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

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *The Modern Language Journal 93*(3), 336–353. DOI: 10.1111/j.1540-4781.2009.00894.x

DeKeyser, R. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*(S1), 1–25. DOI: 10.1111/j.0023-8333.2005.00294.x

Ellis, N.C. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition 38*(2), 341–351. DOI: 10.1017/S027226311600005X

Ellis, N., & Robinson P. (2008). An introduction to cognitive linguistics, second language acquisition, and language instruction. In P. Robinson & N.C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 3–24). NY: Routledge.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Filipović, L., & Hawkins, J.A. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics, 51*(1), 145–176. DOI: 10.1515/ling-2013-0005

Finnish National Agency for Education (2014). *Distribution of lesson hours in basic education.* Retrieved from http://www.oph.fi/english/curricula_and_qualifications/basic_education

Finnish National Board of Education (2014). *Koulutuksen tilastollinen vuosikirja 2014* [Statistical Annuals of Education]. Helsinki: Finnish National Board of Education.

Fraurud, K. (1988). Hierarchical relations between discourse referents and the interpretation of antecedentless definite noun phrases. In V. Rosén (Ed.), *Papers from the tenth Scandinavian Conference of Linguistics* (pp. 196–210). Bergen: University of Bergen.

Genesee, F. (1987). *Learning through two languages. Studies of immersion and bilingual education*. Cambridge: Newbury House Publishers.

Glahn, E., Håkansson, G., Hammarberg, B., Holmen, A., Hvenekilde, A. & Lund, K. (2001). Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition, 23*(3), 389–416. DOI:10.1017/S0272263101003047

Goldschneider, J., & DeKeyser, R. (2005). Explaining the "natural" order of L2 morpheme acquisition in English: A meta-analysis of multiple determinants. *Language Learning, 55*(S1), 27–77. DOI: 10.1111/j.0023-8333.2005.00295.x

Hakulinen A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T.R., & Alho, I. (2004). *Iso suomen kielioppi* [Big Finnish Grammar]. Helsinki: Society of Finnish Literature.

Harley, B. (1993). Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition, 15*(2), 245–260. DOI: 10.1017/S0272263100011980

Harley, B. (1998). The role of focus-on-form tasks in promoting child L2 acquisition. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 156–174). Cambridge, UK: Cambridge University Press.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Cambridge: Newbury House.

Housen, A., & Simoens H. (2016). Introduction: cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38*(2), 163–175. DOI: 10.1017/S0272263116000176

Hyltenstam, K. (1992). Non-native features of near-native speakers. On the ultimate attainment of childhood L2 learners. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: Elsevier Science Publishers.

Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition, 24*(3), 387–418. DOI: https://doi.org/10.1017/S0272263102003029

Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition, 22*(4), 535–556. DOI: https://doi.org/10.1017/s0272263100004034

Järvinen, E-L. (2010). Val av species och bestämdhetsform i ett test för finska grundskoleelever [Choice of definiteness and form of a noun in a test for Finnish pupils in comprehensive school]. In C. Falk, A. Nord & R. Palm (eds): *Svenskans beskrivning 30*. [*Describing Swedish 30*]. Stockholm: Institutionen för nordiska språk, Stockholm University: 154–163.

Laury, R. (1996). Sen kategoriasta – onko suomessa jo artikkeli? [The category of *se* – Does Finnish already have an article?]. *Virittäjä* [*Tuner*], *100*(2), 162–180.

Lieven, E., & Tomasello (2008). Children's first language acquisition. In: P. Robinson & N.C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 168–196). NY: Routledge.

Long, M.H. (1990). The least a second language acquisition theory needs to explain. *TESOL Quarterly, 24*(4), 649–666. DOI: https://doi.org/10.2307/3587113

Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam/Philadelphia: John Benjamins.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*, 37–66.

Nyqvist, E-L. (2013). *Species och artikelbruk i finskspråkiga grundskoleelevers inlärarsvenska. En longitudinell undersökning i årskurserna 7–9*. [Definiteness and use of articles in Swedish of Finnish L2-learners. A longitudinal study from 7th to 9th grade in comprehensive school]. A doctoral dissertation published in electronic format. Turku: University of Turku. http://www.doria.fi/handle/10024/91519

Nyqvist, E-L. (2014). Hurdana species- och artikelfel bör motarbetas? [What kind of definiteness and article errors ought to be avoided?] In J. Lindström, S. Henricson, A. Huhtala, P. Kukkonen, H. Lehti-Eklund, & C. Lindholm (Eds.). *Svenskans Beskrivning 33* [*Describing Swedish 33*] (pp. 353–362). Helsinki: Nordica Helsingiensia, 37.

Nyqvist, E-L. (2015). *Species och artikelbruk hos vuxna svenskinlärare* [Definiteness and article use by adult L2-learners of Swedish]. *NordAnd 10*(2), 75–101.

Nyqvist, E-L. (2016). Swedish as a L3 for Finnish learners: The role of complexity, cross-linguistic influence and study materials in the case of definiteness and the use of articles. *Research in Language, 14*(3), 297–327. DOI: 10.1515/rela-2016-0016

Odlin, T. (2003). Cross-linguistic influence. In C.J. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 436–486). Blackwell Publishing.

Pallotti, G. (2007). An operational definition of the emergence criterion. *Applied Linguistics, 28*(3), 361–382. DOI: 10.1093/applin/amm018

Rahkonen, M., & Håkansson, G. (2008). Production of written L2 Swedish – Processability or input frequencies? In J.-U. Kessler (Ed.), *Processability approaches to second language development and second language learning* (pp. 135–161). Newcastle: Cambridge Scholars Publishing.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Clevedon: Multilingual Matters.

Sajavaara, K. (1983). The article errors of Finnish learners of English. *Umeå Papers in English*, *4*, 72–87.

Schmidt, R.W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 2*(2), 129–158. https://doi-org.ezproxy.utu.fi/10.1093/applin/11.2.129

Sipola, S. (2008). Den språkliga korrektheten i uppsatser på svenska [Grammatical accuracy in narratives in Swedish]. In T. Swanström (Ed.), *Språkbad: Två språk i utveckling. Kielikylpy: Kaksi kieltä kehityksessä* [Immersion: Two languages in development] (pp. 79–91). Turku: Skrifter från Svenska institutionen vid Åbo Akademi.

Teleman, U., Hellberg, S. & Andersson E. (1999a). *Svenska Akademiens Grammatik.* [Grammar of the Swedish Academy] *Vol. 2*. Stockholm: Nordstedts Ordbok.

Teleman, U., Hellberg, S. & Andersson E. (1999b). *Svenska Akademiens Grammatik.* [Grammar of the Swedish Academy] *Vol. 3*. Stockholm: Nordstedts Ordbok.

Trenkic, D. (2007). Variability in L2 article production: Beyond the representational deficit vs. processing constraints debate. *Second Language Research*, 23(3): 289–327. DOI: 10.1177/0267658307077643

Trenkic, D., & Pongpairoj, N. (2013). Referent salience affects second language article use. *Bilingualism: Language and Cognition, 16*(1), 152–166. DOI:10.1017/S1366728912000156.

World Atlas of Language Structures. (2017). *Combination definite articles / indefinite articles.* Retrieved from http://wals.info/combinations/37A_38A#2/25.5/148.5

Vilkuna, M. (1992). *Referenssi ja määräisyys suomenkielisten tekstien tulkinnassa* [Reference and definiteness in Finnish texts]. Helsinki: Society of Finnish Literature.

Definiteness in Written Swedish by Immersion Pupils

Abstrakti

Tutkimuksen tavoite on kahtalainen. Tavoitteena on ensinnäkin analysoida, kuinka 15-vuotiaat, suomenkieliset kielikylpyoppilaat (n=162) hallitsevat ruotsin kielen määräysmuotojen ja artikkelien käytön ja verrata heidän tuotostaan perinteistä ruotsinopetusta saaneiden saman ikäisten (n=67) oppilaiden tuotokseen. Ryhmätason analyysit osoittavat, että kielikylpyoppilaat hallitsevat määräisyyden ja artikkelien käytön tilastollisesti merkitsevästi paremmin kuin kontrolliryhmä, mutta että heillä toisaalta on samantyyppisiä ongelmia kuin aikaisempien tutkimusten tutkimushenkilöillä, joilla ei ole kielikylpytaustaa, kuten epämääräisen artikkelin poisjättöä ja vaikeutta valita substantiivin oikea muoto. Nämä ongelmat ovat kuitenkin harvinaisempia kielikylpyoppilailla. Ryhmätason oppimisjärjestys on hyvin samankaltainen kuin aiemmissa tutkimuksissa havaitut, ja sitä voidaan selittää kompleksisuuden ja kielten välisen siirtovaikutuksen avulla. Yksilötason analyysit puolestaan tuottavat keskenään hieman erilaisia oppimisjärjestyksiä, mikä johtuu analyyseissä käytettyjen analyysitapojen painotuksista.

Abstrakt

Denna studie har två huvudsakliga syften: för det första att analysera hur 15-åriga språkbadselever med finska som L1 (n=162) uttrycker species och använder artiklar i en spontan skrivuppgift och för det andra att jämföra deras produktion med texter skrivna av jämngamla elever i traditionell undervisning (n=67). Analyserna på gruppnivå visar att språkbadseleverna vanligen presterar signifikant bättre än kontrollgruppen, men att de också har likadana problem som informanterna utan språkbadsbakgrund har uppvisat i tidigare undersökningar, såsom bortfall av obestämd artikel och svårigheter med att välja den rätta formen av ett substantiv. Dessa problem är dock mer sällsynta hos språkbadseleverna. Inlärningsgången på gruppnivå påminner om de mönster som upptäckts i tidigare forskning

och som kan förklaras med olika aspekter av komplexitet och tvärspråkligt inflytande. Analyserna på individnivå uppvisar däremot delvis olika inlärningsgångar, vilket beror på de använda kriterierna.

Author's address

Eeva-Liisa Nyqvist

Scandinavian Languages

University of Turku

School of Languages and Translation Studies

Turku Institute of Advanced Studies

20014 Turku

Finland

eesija@utu.fi