The Bottleneck Hypothesis in L2 Acquisition: A Study of L1 Norwegian Speakers’ Knowledge of Syntax and Morphology in L2 English

Isabel Nadine Jensen, Roumyana Slabakova, and Marit Westergaard

1. Introduction

The Bottleneck Hypothesis (Slabakova, 2008; 2013) aims to partially answer the question of why certain properties of language are harder or easier to acquire in a second language (L2). Based on a comparison of different linguistic modules and how they are acquired, the hypothesis holds that functional morphology is the bottleneck and consequently the most challenging part of L2 acquisition. It is argued that the reasons for this are 1) that differences between languages are located in functional morphology (cf. the Borer–Chomsky Conjecture (Baker, 2008; Chomsky, 1995)); and 2) that functional morphology bundles a variety of semantic, syntactic and morphophonological features which affect the acceptability and the meaning of a sentence. Feature Reassembly (Lardiere, 2009) and mismatches in feature realization are responsible for most of the acquisition challenges.

The study presented in this article tests the predictions of the Bottleneck Hypothesis by investigating L1 Norwegian speakers’ knowledge of core syntax and functional morphology in L2 English. It is predicted that Norwegian learners make fewer errors with syntactic operations than with functional morphology, and furthermore, that knowledge of core syntax improves faster than knowledge of functional morphology as the speakers become more advanced.

We test two constructions that do not match in English and Norwegian: subject–verb agreement, obligatory in the L2, and Verb-Second (V2) word order, obligatory in the L1. The former represents knowledge of functional morphology and the latter knowledge of syntax. This is the first experimental study designed to directly test the hypothesis. Our main findings show that participants experienced more challenges with agreement than with verb movement, as they had considerable problems identifying ungrammatical agreement in an acceptability judgment test. This lack of sensitivity was established not only at the lower proficiency levels, but also among the more advanced learners. In addition, there was a stronger correlation between non-V2 word order and proficiency
scores than between agreement and proficiency scores. This fact suggests that learners develop considerably faster in their knowledge of English syntax than in their mastery of functional morphology. We conclude that our findings lend support to the Bottleneck Hypothesis.

2. Theoretical background

Functional morphology, also known as the Functional Lexicon, is identified as the locus of difficulty in L2A for the following reasons: It carries information about grammatical meanings through interpretable features, such as e.g., tense, aspect and definiteness, but also regulates displacement of phrases, i.e., movement, through uninterpretable features. Crucially, functional morphology and all the features it reflects must be learned, just like all other lexical entries.

To exemplify, the features reflected in the morpheme -s on the verb in sentence (1) are [tense], [aspect] and [agreement], but also features that ensure that the subject in English is overt, that it is in the Nominative case, and that the verb stays in the VP, i.e., on the right side of the adverb often. All of these properties have to be acquired for a person to know how to use and interpret this inflectional morpheme.

(1) The chameleon often change-s color.

In a nutshell, the Bottleneck Hypothesis proposes that the functional lexicon is difficult to acquire, whereas universal properties common to all languages are easy to acquire. On this basis, the general predictions of the Bottleneck Hypothesis for L2 acquisition are that functional morphology gives rise to considerable difficulty for L2 learners, especially where features are bundled and there are mismatches between the L1 and the L2; that acquisition of core syntax proceeds with relatively less difficulty; and that universal semantic properties do not create any difficulty for L2 learners.

3. The present study
3.1. Constructions

In this study, we use subject–verb agreement to test learners’ knowledge of functional morphology, whereas word order in declaratives (V2 vs. non-V2) is used to test knowledge of core syntax. As previously mentioned, these are two constructions that exhibit a grammatical contrast in English and Norwegian. The constructions have been chosen on the basis of findings in previous studies, such as Westergaard (2003), who shows that acquisition of English word order is challenging for native speakers of Norwegian at an early stage, as well as Slabakova and Gajdos (2008), who find that subject–verb agreement is problematic not only for beginners, but also for intermediate learners of German L2 (see also Ionin & Wexler, 2002; Haznedar, 2001; Lardiere, 1998a,b). These observations indicate that both verb movement and subject–verb agreement
should be challenging for Norwegian learners of English L2. However, which one of them is more challenging has not yet been investigated. In the following sections, the constructions are discussed in more detail, including the uninterpretable features and their exponents.

English is an SVO language, while Norwegian displays V2 syntax, in that the finite verb is generally required to occur in the second position of main clauses. There are numerous exceptions to V2 in Norwegian, especially with respect to wh-interrogatives across various dialects (see e.g., Westergaard, 2009; Westergaard, Vangsnes & Lohndal, in press) and so-called om-questions (Rognes, 2011). However, in main declarative clauses, the finite verb typically appears in second position (with only a few exceptions related to specific adverbs; see e.g., Westergaard, 2008; Bentzen, 2014). This is seen in both non-subject-initial and subject-initial sentences with adverbs, as illustrated in (2) and (3), respectively.

(2) I går drakk studentene vin.
    Yesterday drank students.DEF wine
    ‘Yesterday the students drank wine.’
(3) Studentene drakk ofte vin.
    Students.DEF drank often wine
    ‘The students often drank wine.’

The V2 word order in Norwegian is caused by V-to-C movement triggered by a syntactic (micro-)cue, an abstract piece of structure in speakers’ I-language grammars (Westergaard, 2009), illustrated in (4).

(4) $\text{Deci}p[\text{XP } \text{Deci}^V]$

The learning task for Norwegian L1 speakers is to unlearn this V2 cue. As Westergaard (2003) shows, Norwegian L1 speakers often transfer this cue to L2 English and unlearning V2 is a difficult process due to the relative infrequency of the relevant contexts in English (non-subject-initial declaratives). In addition, word order is usually not explicitly taught in English L2 classrooms.

Whereas Norwegian has no overt subject–verb agreement morphology, English marks the verb with the suffix –s when the subject is 3rd person singular, illustrated in (5). This is not only frequent in the English input, but also subject to intensive instruction from an early stage in L2 English classrooms.

(5) a. Per og Mari snakker norsk.
      Per and Mari speak Norwegian
      ‘Per and Mari speak Norwegian.’

b. Per snakker norsk.
   Per speaks Norwegian
   ‘Per speaks Norwegian.’

(6) $\text{TP}[\text{XP}_{3sg} \text{T}[V-s]]$
The micro-cue for English subject–verb agreement is illustrated in (6). While L1 children typically acquire this morpheme conservatively (i.e., only provide it where necessary), both under- and overuse has been found in L2A (e.g., Dröschel, 2011; Vettorel, 2014). The learning task in L2 English is to acquire the uninterpretable feature bundle on T (tense, number, person) (Chomsky, 1995, 2000, 2001), in addition to the morphological reflex of the feature bundle, –s, which is a functional morpheme with little salience.

3.2. Research questions and predictions

RQ1: Is morphology more difficult than core syntax in L2 acquisition?
RQ2: Is morphology a more persistent problem than core syntax in L2 acquisition?

To test research questions 1 and 2, we compared the speakers’ knowledge of subject–verb agreement and non-V2 syntax. It was expected that the participants would make more errors with functional morphology and that knowledge of core syntax improves faster than knowledge of functional morphology.

Moreover, previous studies have shown that the difficulty of a construction may depend on factors such as verb type or structural distance. To exemplify, Håkansson and Collberg (1994) found that acquisition of the target word order negation + modal is more difficult than negation + lexical verb in Swedish embedded clauses. In addition, Ocampo (2013) found that subject–verb agreement is more challenging when the structural distance between the agreeing phrases in a sentence is increased, as well as when the number feature on the noun immediately preceding the verb is [plural]. For that reason, morphology was tested by means of four different experimental conditions: long-distance agreement vs. local agreement, and singular vs. plural subjects. Core syntax was tested by varying the verb type in non-subject-initial declaratives, lexical verbs and auxiliaries. This gave rise to a third research question, which is not directly related to the comparison of knowledge of core syntax and functional morphology:

RQ3: Which of the syntactic and morphological conditions are more difficult?

There were three expectations regarding functional morphology: First, it was expected that long-distance agreement would be more difficult than local agreement because sentences with long-distance agreement may cause so called “attraction errors” (see e.g., Bock & Miller, 1991; Bock & Eberhard, 1993). These occur when the subject NP and the NP closest to the verb differ in number. Second, it was expected that long-distance agreement would be more difficult when the subject is singular than when it is plural. These predictions were based on previous research which has found that learners are significantly less sensitive to long-distance agreement errors when the subject is singular and the intervening noun plural (Ocampo, 2013). Finally, regarding local agreement, the prediction was that
the learners would make more errors with sentences that have singular subjects, as most studies have found that the most common agreement error is to drop 3rd person singular –s, rather than to insert a superfluous –s (hypercorrection) (see e.g., Dröschel, 2011).

With respect to the syntactic structures, we predicted that non-V2 syntax would be more challenging when the sentences contain an auxiliary verb. This was based on Pollard’s (1996) argument that auxiliary verbs are unspecified for the feature [inverted] in English, whereas lexical verbs have the feature [-inverted]. That is, learners of English receive input which tells them that the auxiliary may move in some cases in English and that lexical verbs never move out of the VP. This may cause more variability with word order in sentences that contain auxiliary verbs, as learners may be more inclined to move these verbs in English. Nevertheless, both syntactic constructions were expected to cause few problems for the learners, in comparison to agreement.

4. Method
4.1. Participants

Sixty students at two different schools in Norway participated in the experiment. Their ages ranged from 11 to 12 and from 15 to 18. All participants were native speakers of Norwegian, with Norwegian as their only L1, and English as their L2. Their proficiency level in English was measured by means of a subset of a standardized Oxford Proficiency test, consisting of 40 items. Based on this, the participants were divided into four proficiency groups, ranging from low intermediate to advanced.

4.2. Procedure and materials

Data were gathered by means of an acceptability judgement test in which participants were asked to rank sentences on a 1 to 4 Likert scale, with the additional option “I don’t know”, where 1 signified that the sentence was completely unacceptable and 4 marked a completely acceptable sentence. The test was web-based and carried out on Survey Gizmo.

The study took place during school hours, and the participants spent approximately 25–35 minutes to complete the test. The first part was the acceptability judgement test. The participants were presented with five sentences at a time, and they were not able to return to previous pages once they had moved to the next page of the survey. In other words, they were not able to change any of their previous judgements. In addition, the test was pseudo-randomized to ensure that a sentence pair (acceptable and unacceptable) never appeared on the same page or immediately followed one another from one page to the next. Also, pseudo-randomizing made it possible to ensure that there was always at least one sentence that represented syntax, one that represented functional morphology and no more than one filler on each page.
The second part of the experiment was the 40-item proficiency test. This was a multiple-choice task in which the participant had to choose one out of three words or phrases to fill in a blank spot in a sentence. The third and final part of the test was a background questionnaire. Here, the participants were asked to answer questions about their age and linguistic background. This section was in Norwegian to avoid any misunderstandings.

4.3. Experimental conditions and test items

Two experimental conditions were used to test core syntax: Non-subject-initial declaratives with auxiliaries and/or lexical verbs, illustrated in (7)–(8). Four conditions were used to test functional morphology: Long-distance and local agreement in present tense subject-initial declaratives with singular and plural subjects, illustrated in (9)–(12).

(7) Non-subject-initial declaratives with lexical verbs
a. *Yesterday went the teacher to the shop.
b. Yesterday the teacher went to the shop.

(8) Non-subject-initial declaratives with auxiliary verbs
a. *Every day should the students bring their books to school.
b. Every day the students should bring their books to school.

(9) Long-distance agreement with plural subjects
a. *The kids with the red bike plays in the garden.
b. The kids with the red bike play in the garden.

(10) Long-distance agreement with singular subjects
a. *The teacher with black shoes walk to work every day.
b. The teacher with black shoes walks to work every day.

(11) Local agreement with plural subjects
a. *The teachers gives their students a lot of homework.
b. The teachers give their students a lot of homework.

(12) Local agreement with singular subjects
a. *The brown dog play with the yellow football.
b. The brown dog plays with the yellow football.

There were altogether 36 sentence pairs in the acceptability judgement test, with one ungrammatical and one grammatical version of the same sentence, i.e., six sentence pairs in each sub-condition. In addition, there were 13 fillers, which were all ungrammatical. Sentences were of equal length (10–12 syllables), and lexical items were selected from a word frequency list of English to ensure that the sentences were similar in terms of vocabulary. In addition, all subjects were regular full DPs. The non-subject-initial declaratives had past tense lexical verbs and modal auxiliaries to avoid judgments being based on subject–verb agreement. The present tense declaratives that tested agreement had no irregular verbs, and the ones that tested long-distance agreement all contained prepositional phrases embedded in the subject DP. There was also an asymmetry between the
number of the head noun and the number of the noun that immediately preceded the verb in these sentences.

5. Results
5.1. The proficiency test

As explained in section 4.1, the participants’ proficiency levels were determined on the basis of a 40-item subset of a standardized Oxford Proficiency test. The participants received one point for each correct answer, so that the highest possible score was 40. Based on this test, participants with a score lower than 10 were considered beginners; a score between 10 and 32 indicated an intermediate level, and a score above 32 classified learners as advanced. In addition, the intermediate level was split up into low intermediate, intermediate and high intermediate.

In this experiment, the participants’ scores ranged between 11 and 38 points. This resulted in the following proficiency groups: low intermediate learners (between 11 and 17 points; N = 9), intermediate learners (between 18 and 22 points; N = 11), high intermediate learners (between 26 and 32 points; N = 20) and advanced learners (between 33 and 38 points; N = 20). The number of participants in each group was not decided beforehand. As there was a strong positive correlation between age and proficiency scores (adjusted $r^2 = .7753$), only proficiency scores were used as a variable in the analysis of the data.

5.2. The acceptability judgement test

The p-value was set to .05, which means that any value smaller than this number is statistically significant. T-tests and regressions were used to analyze the data. The Likert scale was treated as a binary variable with the categories unacceptable (values 1 and 2) and acceptable (values 3 and 4).

5.2.1. Agreement vs. non-V2 syntax

Figures 1 and 2 display the mean scores of each proficiency group. Figure 1 shows that the mean acceptability scores on agreement was relatively high for both grammatical and ungrammatical agreement at all proficiency levels, i.e., all groups accepted most grammatical sentences. Ungrammatical agreement was also frequently accepted, well into the advanced stage. That is, even at the highest proficiency level, the mean acceptability score for ungrammatical sentences was relatively high.
Figure 1: Mean acceptability score on agreement

Figure 2 illustrates the mean acceptability scores for verb movement. The low intermediate and intermediate speakers tended to accept both ungrammatical and grammatical V2 syntax, i.e., there were few errors with grammatical sentences and many errors with ungrammatical sentences. At the high intermediate and advanced levels, on the other hand, the participants made few errors with both ungrammatical and grammatical sentences, as they tended to accept non-V2 and reject V2 word order. This is different from these two groups’ score on agreement.

Figure 2: Mean acceptability score on verb movement
5.2.2. Grammatical vs. ungrammatical sentences

There was a weak correlation between proficiency and acceptability scores on grammatical sentences in both conditions. That is, the adjusted $r^2$ value was .1159 for the correlation between proficiency scores and non-V2 syntax and .06865 between proficiency and agreement. The mean scores were all high, indicating that the participants made few incorrect judgements on all grammatical sentences.

With ungrammatical sentences, there was a strong negative correlation between proficiency scores and correct judgements of non-V2 syntax (adjusted $r^2$ = .4688). This indicates that participants with higher proficiency scores rejected ungrammatical verb movement more frequently, whereas participants with lower proficiency scores often accepted these sentences. With ungrammatical agreement, on the other hand, the negative correlation was relatively weak (adjusted $r^2$ = .2227), which indicates that a larger group of participants with higher levels of proficiency often accepted ungrammatical agreement. In other words, there was more variability on agreement than could be explained by proficiency.

When comparing ungrammatical verb movement and ungrammatical agreement, we find that there were significantly more errors with the latter condition in each of the proficiency group; i.e., in each group, the participants accepted significantly more ungrammatical agreement than ungrammatical verb movement. This is illustrated in Figure 4, which plots accuracy in rejecting ungrammatical test sentences across the four proficiency groups.

![Figure 3: Accuracy in rejecting ungrammatical verb movement and agreement](image)

5.2.3. Sub-conditions

In the low intermediate and intermediate groups, there were no statistically significant differences between the acceptability scores on the sub-conditions for
subject–verb agreement. With respect to grammatical subject–verb agreement, the high intermediate and advanced groups made significantly more errors with long-distance agreement with plural subjects in comparison to each of the other morphological sub-conditions ($p < .05$). Moreover, the advanced group made significantly more errors with long-distance agreement than local agreement with singular subjects. This suggests that long-distance agreement with plural subjects was the most problematic morphological sub-condition.

In addition, the advanced speakers rejected grammatical local agreement with plural subjects significantly more often than with singular subjects. Similarly, they accepted ungrammatical local agreement with plural subjects significantly more than with singular subjects. In other words, the advanced speakers made more errors with local agreement when the subjects were plural, both in grammatical and ungrammatical sentences. This implies that also with local agreement, sentences with plural subjects were more problematic than sentences with singular subjects.

Regarding the syntactic sub-conditions, there were no statistically significant differences in the low intermediate group. In the intermediate, high intermediate and advanced groups, on the other hand, the speakers make significantly more errors with ungrammatical verb movement when the sentences contained an auxiliary verb. In other words, they tended to accept these types of sentences more often than ungrammatical sentences with lexical verbs.

Since long-distance agreement is clearly the most challenging sub-condition and may be caused by so-called attraction errors (see section 3.2.), the dataset was also analyzed without this construction. The correlation between proficiency scores and judgements was still relatively weak for both grammatical (adjusted $r^2 = .2524$) and ungrammatical agreement (adjusted $r^2 = .2062$). In other words, most participants accepted both grammatical and ungrammatical agreement, and excluding long-distance agreement only had a minor effect on the results.

6. Discussion

6.1. Is functional morphology more difficult than core syntax?

The expectation regarding research question 1 was that subject–verb agreement would be more difficult than non-V2 syntax (see section 3.2.). As illustrated in section 5, the mean acceptability score on ungrammatical agreement was significantly higher than the score on ungrammatical verb movement for all proficiency groups, showing that there was a higher acceptance of ungrammatical agreement than of ungrammatical verb movement. In contrast, each group made fewer errors in both conditions when the sentences were grammatical.

The fact that there were fewer errors with grammatical sentences does not necessarily mean that the participants’ performance is good. In fact, the low number of errors with subject–verb agreement indicates weak performance skills on this condition, as the participants were not able to distinguish between acceptable and unacceptable sentences. In contrast, the participants rejected ungrammatical verb movement and accepted grammatical non-V2 syntax with a
higher accuracy, which suggests good performance on verb movement. The same results were found when long-distance agreement was excluded from the dataset.

These observations reveal that unlearning the V2 cue may be less problematic than learning the rules of subject–verb agreement in English L2 acquisition. Such findings lend support to prediction 1: acquiring subject–verb agreement would be more difficult than unlearning verb movement. Furthermore, our results support the previously mentioned findings in Haznedar (2001), Ionin and Wexler (2001) and Lardiere (1998a,b) that L2 speakers of English are more accurate with syntactic phenomena in obligatory contexts than with morphosyntactic phenomena related to the same functional category, such as subject–verb agreement and verb movement. As White (2003) argues, such results suggest that L2 learners are able to engage knowledge of syntactic operations even at a time when they do not produce the target functional morphology, a position she dubs syntax-before-morphology. The opposite view, morphology-before-syntax, claims that acquisition of functional morphology drives acquisition of functional categories. Our findings offer support for the former position. In addition, this is the first study showing that the discrepancy previously attested in production is also attested in comprehension, as measured by acceptability judgments.

6.2. Is functional morphology a more persistent problem than core syntax?

In order to test the Bottleneck Hypothesis, it is necessary to not only investigate the difficulty of agreement and non-V2 syntax, but also the way in which acquisition of these two properties develops. As shown by the results in section 5.2.2., there is a weak negative correlation between proficiency and acceptability scores on ungrammatical agreement, which suggests that although proficiency in English increases, ungrammatical agreement is still frequently accepted. With verb movement, on the other hand, the correlation was strong, which indicates that when learners become more advanced, they are able to accurately reject ungrammatical verb movement.

With respect to the proficiency groups, the general tendency was that the participants did not recognize ungrammatical agreement until the advanced stage, but even at this level the acceptability score was relatively high (46.7). In contrast, recognizing ungrammatical verb movement seemed to be unproblematic at the high intermediate level. In other words, performance on verb movement improved considerably from the intermediate to the high intermediate stage, whereas ungrammatical agreement seemed to stagnate. This supports prediction 2: subject–verb agreement is a more persistent problem than verb movement in L2 acquisition of English.

6.3. Which of the syntactic and morphological conditions are more difficult?

With respect to the morphological conditions, the results in section 5.2.4. show that long-distance agreement is more difficult than local agreement, i.e., that adding an additional element between the subject and the verb causes problems
for the learners. This adds to findings in previous studies, such as Ocampo’s (2013) comprehension study, which shows that more structural complexity in a sentence makes agreement more challenging.

In addition, subject–verb agreement was more difficult with plural than with singular subjects for our participants, and that was the case in sentences with both local and long-distance agreement. That is, participants were more accurate when the morpheme -s was needed. This result was not expected, as previous studies have shown that most errors in both comprehension and production studies occur when the subject is singular, both in sentences with local and with long-distance agreement (see section 3.2). In other words, the results found in the current study go against the plural markedness effect, which holds that the marked feature [plural] on the noun immediately preceding the verb disrupts the establishment of agreement (see Ocampo, 2013). In general, high variability in evaluating sentences with morphological errors lends support to the Missing Surface Inflection Hypothesis (Haznedar & Schwartz, 1997; Prevost & White, 2000) because it suggests deteriorated lexical access to morphemes under conditions taxing working memory.

Regarding the syntactic conditions, sentences with auxiliary verbs were more problematic than sentences with lexical verbs, as expected. This supports the argument that learners are more inclined to move an auxiliary verb in L2 English. Auxiliaries do move to the CP-domain in some contexts in English (in questions), whereas lexical verbs always stay in the VP. This may cause learners to be less sensitive to auxiliary verb movement in non-subject-initial declaratives, especially as they are infrequent in the input (see section 3.2).

6.4. Is the Bottleneck Hypothesis supported?

We argue that the results presented in section 5 lend support to the Bottleneck Hypothesis, as speakers’ performance with functional morphology was weaker than with syntactic movement, as measured by accuracy at a point in time. Furthermore, performance with functional morphology lingered at a low accuracy level, as measured in development across groups increasing in general proficiency. We argue that the reason for this is that the V2 property has a syntactic exponent (verb movement), while agreement has a morphological exponent. In addition, the morphological exponent of agreement subsumes other interpretable and uninterpretable features reflected in the same functional category Tense, but it is not crucial in communication, since the subject is overt in English.

7. Conclusion and suggestions for future research

The purpose of the study presented here is to contribute to current knowledge of the cognitive process of L2 acquisition by testing the predictions put forward by the Bottleneck Hypothesis. We have carried out an experiment on the L2 acquisition of English, focusing on learners’ knowledge, and development of knowledge, of functional morphology and core syntax. In sum, although there are
remaining questions, the experiment supports the Bottleneck Hypothesis. Furthermore, the results have implications for language teaching, as they illustrate that even at advanced proficiency levels, learners struggle with subject–verb agreement, although this property is explicitly taught in Norwegian schools. In order to further investigate the hypothesis, it is necessary to also look at functional morphology in comparison to other domains than core syntax, such as semantics or the interfaces, as well as different language combinations. It is also essential to use on-line experiments, such as eye-tracking or speeded acceptability judgement tests, in order to investigate L2 learners’ more automatic and implicit knowledge.

References


