Running head: AGREEMENT INFLECTION IN CHILD L2 DUTCH

The impact of verb form, sentence position, home language and L2 proficiency on subject-verb agreement in child L2 Dutch

Elma Blom[1,2] and Harald R. Baayen[2]

[1] University of Amsterdam

[2] University of Alberta

This research was supported by a Marie Curie International Fellowship with the 7th European Community Framework Programme (IOF-219276 to Elma Blom) and by the Dutch Netherlands Organisation for Scientific Research (NWO #360-70-110 to Hans Bennis and Fred Weerman). We are grateful to Leonie IJzereef and Jitka Horáková for collecting data, transcribing and coding. Antje Orgassa and Daniela Polišenská helped designing the experiment and developing the materials.

Correspondence concerning this article should be addressed to Elma Blom, 4-32 Assiniboia Hall, University of Alberta, Edmonton, Alberta, Canada T6G 2E7. Email: w.b.t.blom@uva.nl

Abstract

It has been argued that children learning a second language (L2) omit agreement inflection due to communication demands (Haznedar, 2001; Haznedar & Schwartz, 1997; Ionin & Wexler, 2002; Prévost, 2003). The conclusion of these studies is that L2 children know the morphological and syntactic properties of agreement inflection, but sometimes insert an inflectional default form (i.e. the bare verb) in production. The present study focuses on factors that explain errors with subject-verb agreement in the speech of children learning Dutch as their L2. Analyses of experimentally obtained production data from four to nine year old L2 children reveal that verb form, sentence position, home language and L2 proficiency determine accuracy with subject-verb agreement in the L2. Most errors were omissions of inflection, in line with the above hypothesis. However, in more exceptional contexts, the children also substituted verb forms, which is more difficult to reconcile with the claim that L2 children's errors reflect insertion of a default form. [159 words]

The impact of verb form, sentence position, home language and L2 proficiency on subject-verb agreement in child L2 Dutch

It is a well-established observation that second language (L2) learners frequently err when using grammatical morphemes (Bailey, Madden, & Krashen, 1973; Dulay & Burt, 1973, 1974; Goldschneider & DeKeyser, 2001; Larsen-Freeman, 1975, 1976). Various recent studies have highlighted problems with tense and agreement morphemes in child L2 development (Haznedar, 2001; Haznedar & Schwartz, 1997; Herschensohn, Stevenson, & Waltmunson, 2005; Ionin & Wexler, 2002; Prévost, 2003; Paradis, Rice, Crago, & Marquis, 2008). The conclusion drawn in these studies is that L2 children know the morphological and syntactic properties of these morphemes, but nevertheless fail to use them in production.

Little is known about the factors that cause L2 children's errors with verb inflection in speech production. Therefore, in this study, we examined the contribution of multiple factors on L2 Dutch children's production of agreement inflection. The use of elicitation techniques allowed the collection of information across a range of syntactic contexts, that varied in person, number and word order. The children in this study were between ages four and nine, and had either a richly inflecting or an isolating home language. The outcomes of this study indicate that verb form, sentence position, home language and L2 proficiency impact on L2 children's accuracy with subject-verb agreement in speech production.

Verb form

The sentences in (1) and (2) illustrate examples from a Turkish L1 child learning L2 English and an English L1 child learning L2 German, who both fail to express the relation

between subject and verb. In (1) the child has omitted the English third person singular (3SG) suffix *-s*. In (2), the German 3SG suffix *-t* is left unexpressed.

(1)	Dinosaur	turn	back	and	drink	water	(Haznedar, 2001, p. 26)
	Dinosaur	turn-ø	back	and	drink-ø	water	
	Target:	turns					

(2)	Mein	Bruder	heiss	Marc	(Prévost, 2003, p. 78)
	My	brother	call-ø	Marc	
	Target:		heisst		

In order to explain omission errors such as in (1) and (2), researchers have proposed the Missing Surface Inflection Hypothesis or MSIH (Haznedar, 2001; Haznedar & Schwartz, 1997; Ionin & Wexler, 2002; Prévost, 2003; Paradis et al., 2008). This hypothesis holds that omission of inflection in child L2 English and German results from L2 children's reliance on a morphological default form, a zero morpheme, which is less specified than other forms in the paradigm, and is inserted if lexical access to more specific verb forms is hindered or delayed.

On the MSIH, grammatical morphemes are mapped to functional features which are derived from a universal set of features (Lardiere, 2009). For instance, the inflectional morphemes *–s* (in English) and *–t* (in German) are associated with finiteness, person (3) and number (SG). Discussing English, Ionin and Wexler (2002, p. 118) state that uninflected bare verb forms "are finite forms that have simply not received morphological agreement marking

and that can be used across all tense/agreement contexts."[1] With respect to German, Prévost (2003, p. 88) assumes that bare forms "are unspecified for person features but specified as [+finite]". According to the theoretical framework assumed in the aforementioned studies underspecified forms are default forms.[2] Access to this default is in normal conditions blocked by inflected forms that are associated with person and number features. If for some reason retrieval of the correct inflected form fails, the default rule applies, resulting in incorrect bare verbs as in (1) and (2).

The MSIH predicts that L2 English and L2 German children omit inflection, and do not substitute inflected verb forms. A predominance of omission errors has indeed been reported for child L2 English (Ionin & Wexler, 2002). English has a very poor inflectional paradigm with many contexts where the bare verb form is grammatical. Hence, L2 English children's overuse of this form may, to a certain extent, be expected. Chilla (2008) describes the development of subject-verb agreement in four typically developing L2 German children with L1 Turkish. German has a richer agreement paradigm than English with four different verb forms in the present tense indicative paradigm and provides more possibilities for substitutions. Two of the L2 German children in Chilla's study showed the omission profile. These children started to acquire German around the age of three. Two other children, who had an age of onset around age six, also substituted inflected forms, e.g., 2SG *–st* in 3SG contexts or 3SG *–t* in plural contexts. The differences between children may suggest that a

---

[1] A detailed version of the same analysis of English is provided by Wexler, Schaeffer, & Bol's (2004) study of children with specific language impairment.

[2] This is the framework of Distributed Morphology (Halle & Marantz, 1993; Halle, 1997; Harley & Noyer, 2001).

later starting age contributes to the variability in errors with subject-verb agreement (see also Meisel, 2009).

Herschensohn, Stevenson and Waltmunson (2005) investigated errors with subject-verb agreement in the data of L2 Spanish in six- and seven-year old Anglophone children. Spanish is a pro-drop language that marks both person and number through inflection on the verb. The present tense indicative paradigm consists of six different froms; there are three verb classes. In this study, the children's age of onset of exposure to Spanish was five. The children participated in a picture description task that elicited lexical verbs in 3SG and 3PL contexts. The results of Herschensohn et al.'s study are interesting for two reasons. First, in this study, the largest percentage of errors (87%) involved substitution of inflection rather than missing inflection. Secondly, the most frequent error was singular/plural reversal: 3SG forms in 3PL contexts and 3PL forms in 3SG contexts. Both errors occurred approximately equally often.

The different studies indicate that the MSIH may not be equally relevant across languages. Therefore, in this study we investigated another target language: Dutch. Like English, Dutch has grammatical and ungrammatical contexts for the bare verb, but agreement marking in Dutch is more complex than agreement in English.

Sentence position

According to the MSIH, inflection drop occurs as the effect of failure to block the default rule. It has been suggested that "retrieval and communication pressures may cause learners to sometimes leave out inflectional morphemes in production." (Ionin & Wexler, 2002, p. 128). In other words: Because more specific verb forms cannot be retrieved on time, bare verbs are used. Studies that looked more closely at the role of production consistently

report that L2 children who make errors with tense and agreement inflection in production are able to comprehend these morphemes (Herschensohn et al., 2005) or judge whether omission of these morphemes is grammatical or not (Ionin & Wexler, 2002; Paradis et al., 2008). Such asymmetries may indeed point to performance factors related to speech production. However, strikingly little is known about which performance factors impact L2 children's omissions of inflection in production.

Given the findings in a recent study on L1 English development by Song, Sundara and Demuth (2009), a verb's position in the sentence could be one relevant factor. According to Song et al., children drop 3SG *-s* inflection more frequently in medial position than in sentence final position because children may have more time to produce coda consonants and morphemes phrase-finally, as compared to phrase-medially. In addition, producing 3SG *-s* in final position may be less challenging from the perspective of articulation/planning than producing this morpheme in medial position, because no word is following 3SG *-s* in this position. In medial position, the immediate occurrence of another word may neccessitate additional articulatory gestures. Song et al. report naturalistic and experimental data from children between ages 1;3 and 3;6 showing that 3SG *-s* is more often omitted in medial position than in sentence-final position. This effect was amplified with increasing sentence length. A similar effect of sentence position was found in Dalal and Loeb's (2005) study on L1 English children diagnosed with SLI: 5-year olds with SLI were more likely to produce past tense *–ed* in final position than in the middle of a sentence.

Both factors mentioned by Song et al. (time and articulation/planning) are performance factors that may impact speech production. These factors can apply to language users irrespective of age and type of acquisition, and hence might also explain why L2 children sometimes drop agreement inflection in production. Therefore, the second goal of

our study is to explore the effect of sentence position on L2 Dutch children's ability to produce agreement inflection in obligatory contexts.

Transfer

Unlike monolingual children, L2 children have prior knowledge of another language which can shape their performance with inflection in the target language. The question whether L2 children transfer L1 properties in the domain of inflection is of theoretical interest because different hypotheses have been proposed that do not make the same predictions. A study of child L2 acquisition can enhance our insight into effects of transfer because theories of transfer are generally developed for adult L2 acquisition and may not necessarily apply to L2 children (Unsworth & Blom, 2010). Below we expand on these issues.

According to the Full Access Full Transfer hypothesis, (adult) L2 learners transfer all morpho-syntactic information from their L1 to their L2 (Schwartz & Sprouse, 1996; White, 2003).[3] If this hypothesis would apply to child L2 acquisition, we expect that children with richly inflecting L1s are in general more accurate with L2 inflection than children with isolating L1s. However, it has been suggested that L2 children's development in the domain of inflection may be more similar to the development of L1 children than to the development of L2 adults (Schwartz, 2004). Also, researchers have proposed that L2 learners, irrespective

[3] A more specific instantiation of the Full Access Full Transfer hypothesis is the Prosodic Transfer Hypothesis, which has also been proposed to explain drop of grammatical morphemes in adult L2 acquisition (Goad & White, 2004, 2008). For the purpose of the present study, we did not distinguish between specific interpretations of the Full Access Full Transfer hypothesis.

of age of onset, may transfer word order properties from their L1 but will not transfer inflectional properties (Vainikka & Young-Scholten, 2010). Such hypotheses would predict that the child L2 acquisition of inflection is not impacted by inflectional properties of the L1.

Analyses of longitudinal naturalistic data indicate that L2 English children whose L1 has rich agreement reach 90% correct use of 3SG *-s* in obligatory contexts after fewer months of exposure than L2 English children with isolating L1s; this observation suggests that L2 children transfer L1 properties in the domain of inflection (Blom, Paradis, & Sorenson-Duncan, in press). L2 English children with inflecting L1s also perform better than children with isolating L1s with respect to (a composite score of) 3SG *-s* and past tense *-ed* in elicited production experiments (Paradis, in press). After a long period of exposure child L2 learners of English with an inflecting L1 (Spanish) are as accurate as native speakers with respect to judging sentences that involve errors with English inflectional morphology, whereas child L2 learners of English with an isolating L1 are less accurate (McDonald, 2000). However, Paradis (2005) and Paradis et al. (2008) found little evidence for effects of L2 English children's L1 on 3SG *-s* and past tense *-ed*.

The contrasting conclusions of these studies underscore the importance of more research in this area. Therefore, in this study, L2 children have been assigned to two L1 groups based on inflectional properties of their home language.

Age and L2 proficiency

Previous studies have indicated that older L2 children perform better due to more L2 experience (Gathercole, 2002a, 2002b, 2002c) and/or greater cognitive maturity (Gathercole, 2007; Golberg, Paradis, & Crago, 2008; Paradis, in press). In the present study, both chronological age and L2 proficiency were included. The L2 proficiency measure in our

study may index language aptitude, which is the prediction of how well, relative to other individuals, an individual can learn a foreign language in a given amount of time and under given conditions (Carrol & Sapon, 2002). Language aptitude is a consistent predictor of success in adult L2 acquisition (Dörnyei & Skehan, 2003), and it may be expected to have a similar role in child L2 acquisition. L2 proficiency may also be relevant in relation to effects of transfer, because effects of transfer are not expected to influence L2 children's development throughout. This is confirmed by a recent study of article acquisition in child L2 English (Zdorenko & Paradis, 2008). Zdorenko and Paradis compared L2 English children with [+article] and [-article] L1s. Children in the latter group dropped articles more often than children in the former group, but this difference only emerged in the first round of data collection; hence in a very early developmental stage when the children had lower L2 proficiency.

Aims of the study

The aim of the study is to investigate factors that determine L2 children's accuracy with subject-verb agreement in speech production. In order to evaluate hypotheses concerning the role of verb form, sentence position, home language, age and L2 proficiency on L2 children's production of subject-verb agreement, the following four research questions were formulated:

(i) To what extent does verb form - bare or inflected - impact L2 children's accuracy with subject-verb agreement?

(ii) Does sentence position - medial versus final - influence L2 children's accuracy with subject-verb agreement?

(iii) Is the development of subject-verb agreement determined by inflectional properties of L2 children's home language?

(iv) To what extent is the development of subject-verb agreement influenced by L2 children's age and their proficiency in the L2?

Questions (i)-(iv) will be addressed through analyses of data obtained in two elicitation tasks. Below we explain the relevant properties of Dutch and design of the study in relation to questions (i)-(iv).

Dutch is a Subject-Object-Verb (SOV) language. In main clauses, the finite verb (which expresses tense and agreement) follows the first constituent due to a Verb Second (V2) rule; this results in SVO orders. In 1SG contexts, the bare verb stem is used, in 2SG and 3SG contexts, stem+t appears and in plural contexts, stem+en is the required verb form. Examples are given in (3)-(5):

(3) Ik wandel vaak.
I walk-1SG often

(4) Jij/hij/zij wandelt vaak.
You/he/she walk-2/3SG often

(5) Wij/jullie/zij wandelen vaak.
We/you/they walk-PL often

For purposes of emphasis, a constituent can be fronted. In this case, subject and finite verb shift places so that V2 is maintained (inversion) (e.g. OVS in case an object is fronted, AdvVS in case an adverb is in first position). In such clauses, 2SG subjects are preceded by the bare verb, as illustrated in (6).

(6) Morgen wandel jij.
Tomorrow walk you-2SG

Bare verbs are grammatical in 1SG contexts, irrespective of word order. However, the use of bare verbs in 2SG contexts is marked. Therefore, 2SG sentences with subject-verb inversion would typically be a context where substitution errors occur. Other inflected forms in inversion sentences are identical to those of subject-verb agreement in SVO sentences. In verb-final structures (SOV), such as subordinate clauses, subject-verb agreement is similar to SVO sentences. The comparison between SVO and SOV orders is relevant for assessing the impact of medial versus final sentence position.

Using the experimental method of elicitation, we systematically collected information on L2 children's performance across syntactic contexts and word orders. An experimental

procedure was chosen because in naturalistic data the chances of obtaining data from a group of children that are comparable (e.g., use of inflection with the same verbs, contexts and word orders) would be very low (Eisenbeiss, 2010).

The outcomes of the experiment were organized in three separate datasets; this organization allowed addressing the first two research questions through appropriate statistical analyses. Table 1 gives an overview of the first dataset, with the target forms between brackets. Dataset 1 contained data on 1SG, 2SG, 3SG, 1PL and 3PL contexts in sentences with the unmarked SVO order; in Table 1 these are the cells within the area marked by the dotted lines. The second dataset comprised data on 2SG and 3SG contexts in SVO sentences and main clauses with subject-verb inversion and an adverb in first position (AdvVS); dataset 2 is marked by the grey cells in Table 1. The third set of data included 3SG contexts in SVO and SOV sentences; in Table 1 these are the cells in the area marked by the black line.

#INSERT TABLE 1

Datasets 1 and 2 are relevant for the first research question: To what extent does verb form - bare or inflected - impact L2 children's accuracy with subject-verb agreement? Given the MSIH, we expect higher accuracies in contexts that target bare verbs than with contexts targeting inflected verbs. Errors with inflected verbs are expected to be errors of omission. The third dataset is relevant for the second research question: Does sentence position - medial versus final - influence L2 children's accuracy with subject-verb agreement? It is expected that more errors are made in SVO than in SOV sentences. In particular do we expect more drop of 3SG *-t* in SVO sentences than in SOV sentences.

For investigating the third research question - Is the development of subject-verb agreement determined by inflectional properties of L2 children's home language? -, we assigned the L2 children to two L1 groups based on inflectional properties of the children's home languages (which was also their L1). Children in one group had a richly inflecting L1 with both person and number agreement, whereas children in the other group had an isolating L1. If transfer impacts L2 children's development of verb inflection, as predicted by the Full Access Full Transfer hypothesis, we expect to find differences between these groups in how accurate they are with L2 subject-verb agreement.

With respect to the fourth research question - To what extent is the development of subject-verb agreement influenced by L2 children's age and their proficiency in the L2?-, we expect older L2 children to perform better than younger children and that more proficient L2 children make fewer errors with subject-verb agreement than less proficient L2 children. Finally, it is expected that differences between L1 groups are more prominent in lower-proficiency L2 children than in higher-proficiency L2 children.

Method

Participants

Sixty-four children were tested. Two children were excluded post-hoc because of their low number of analyzable responses. The children came from families that immigrated from China, Morocco and Turkey and now lived in three major cities in the western part of the Netherlands (Amsterdam, Rotterdam, Utrecht). Children with apparent language or speaking disorders, or who were enrolled in speech therapy were excluded from the study.

Selected children communicated with their caregivers in Mandarin, Cantonese, Moroccan-Arabic (sometimes in combination with Tarifit, a Berber language) or Turkish.

Mandarin and Cantonese are tonal languages that do not inflect for person, number or gender. Another property of both Mandarin and Cantonese that may be relevant with respect to suffixation processes is that these languages do not allow consonant clusters in word-final position. Mandarin is more restricting than Cantonese and allows only the consonants [n] and [ŋ] in the coda; [r] can be attached as a suffix. Cantonese also allows the stops [p], [t] and [k] in word-final position (Lin, 2001; Matthews & Yip, 1991). Participants in a second group spoke Moroccan-Arabic (sometimes in combination with the Berber language Tarifit) or Turkish at home. Moroccan-Arabic and Tarifit differ in lexicon and morphophonology, but are rather similar with respect to subject-verb agreement. Both languages have rich agreement paradigms consisting of seven or eight different forms with distinctions between person (1, 2, 3), number (SG, PL), and between feminine and masculine in 3SG (perfective) or 2/3SG (imperfective). In both languages, word-final consonants and consonant clusters are allowed (Hoogland, 1996). Turkish is an agglutinative, richly inflecting language. The Turkish present tense indicative paradigm has six different forms that distinguish between person (1, 2, 3) and number (SG, PL). Turkish also allows consonants in word-final position as well as consonant clusters (Lewis, 1967).

The home language was the L1 of the children, who had limited active and passive knowledge of Dutch when they began attending school at age four. At the age of four, the children were either unbalanced bilinguals with Dutch as their (much) weaker language or sequential bilingual children, even though nearly all children were born in the Netherlands. Information about a child's level of Dutch when (s)he started attending school was gathered during an interview with the teacher using an adapted version of the *Anamnese Meertaligheid* 'Questionnaire on the child's multilingual situation' (Blumenthal & Julien, 2000). In schools,

Dutch was the official language, hence going to school led to a significant change in the children's exposure to Dutch.

L2 proficiency was based on children's performance on the sentence repetition task of the *Taaltoets Alle Kinderen (TAK)* (Verhoeven & Vermeer, 2002). The TAK is a standardized test procedure to assess oral proficiency in Dutch native and non-native children. The sentence repetition test consisted of twenty sentences, each of which was scored with respect to the child's realization of one specific word order property and a function word. The targeted properties did not include subject-verb agreement. Two points were assigned if both the specific word order and the function word were repeated correctly. If only one of the two was correctly repeated, one point was assigned, and if neither was correctly repeated, zero points were assigned.

Table 2 gives an overview of properties of the participants, distinguished by their country of descent:

INSERT TABLE 2

Materials

Both tasks targeted the same three verbs (drinking, cleaning, drawing), which were imageable and transitive. The relevant activities were represented by high-quality full-colour photos. The photos showed no distracting visual information and focused on the activity (see Appendix for examples of items used in both the picture description and in the situation description task). All items were presented in a pseudo-randomized order in order to prevent immediate occurrence of the same verb after each other. For each task, experimental items were preceded by practice items introducing each experimental condition: Five items for the

picture description task and three items for the situation description task. For practicing, images were used that depicted the actions of reading and cutting.

The experimental items in the picture description task were presented in a booklet mixed with items testing choice of definite determiners and use of pre-nominal adjectives (Blom, Polišenská, & Weerman, 2008). In the 2SG and 3SG AdvVS condition, two pictures depicting the same activity but that differed in actor were placed next to each other. In the 3SG and 3PL SVO condition and the 3SG SOV condition, the two adjacent pictures did not differ in person/number of the actor but the actor(s) manipulated a different object. The number of experimental items was 24. In the situation description task cards were used that depicted actions that were then acted out by both the participant and the experimenter. The number of experimental items was nine.[4]

Procedure

In the picture description task the participant was instructed to describe a difference between two adjacent pictures by completing a sentence initiated by the experimenter. In this context Dutch uses present tense indicative forms. In the main clause conditions, the experimenter elicited a response by pronouncing the subject (referring to the actor(s) on the pictures). In the picture description task, SVO conditions had two subconditions: 3SG (*De man [drinkt koffie*] 'The man is drinking coffee'; the target response is between square brackets) and 3PL (*De meisjes [drinken sap]* 'The girls are drinking juice'). The SOV condition had no subconditions; the experimenter's lead-in required a relative clause with a 3SG verb as response (*Dit is de man die [koffie drinkt]* 'This is the man who is drinking

[4] Materials, guidelines for testing procedures and coding conventions can be obtained through contacting the first author (Blom, Orgassa, & Polišenská, 2008).

coffee'). Both in the SVO and SOV conditions the two adjacent pictures did not differ in person/number properties of the actor, but the crucial difference was between the object manipulated in the two pictures. This method forced the participants to use a direct object, which was needed to determine whether the correct word order was used, which, in turn, was relevant for assessing the impact of sentence position. AdvVS conditions had two subconditions: 3SG (*Daar [drinkt hij/de man]* 'There he/the man is drinking') and 2SG (*Hier [drink jij]* 'Here you are drinking'). In these conditions, the experimenter pronounced a sentence-initial adverb. Adjacent pictures differed in person/number properties associated with the actor, namely 2SG versus 3SG. Items testing 2SG agreement depicted the experimenter.

1SG, 2SG and 1PL contexts were tested with a situation description task, which had the form of a game. This different task was chosen to create a natural context for subject-verb agreement with discourse-related pronouns. The experimenter as well as the participant simultaneously picked up a card from two strictly ordered piles and turned it around. Each card showed an image of an ongoing action. After seeing the image, experimenter and participant performed the action with the help of attributes (e.g. a brush, a glass, a book) which lay before them on the table. The participant's task was to describe the ongoing situation. There were two situations possible, depending on whether experimenter and participant picked up similar or different cards: Experimenter and participant either performed the same action or they performed different actions. In the first situation the target response contained a verb in 1PL form (*We drinken koffie* 'We are drinking coffee'). In the second situation the target response contained a verb in 1SG (*Ik drink koffie* 'I am drinking coffee') and 2SG form (*Jij tekent een boom* 'You are drawing a tree').

Test sessions took place at the children's schools in a separate room. Each participant was tested individually. The test sessions were recorded on a minidisc, and transcribed afterwards and coded for the categories described in the next section.

Coding and analyses

All verbs that appropriately described the pictures were accepted as scorable responses. Responses with auxiliary verbs were excluded from the analyses because Dutch auxiliaries have irregular agreement paradigms; 14% of the responses were excluded for this reason (see, for more information on the use of auxiliary verbs in L2 Dutch children in a similar task: Blom & De Korte, 2011). Other excluded responses comprised unintelligible responses, responses in which the participant used a verb that did not pertain to the activity depicted by the picture, responses with incorrect word order and responses that did not contain sufficient information on the verb's position (18%). As explained above, in the SVO and SOV conditions responses had to contain an object in order to determine a verb's position in the sentence (that is, medial or final); therefore, in these conditions SV responses were excluded. In the AdvVS condition, responses had to contain a subject in order to determine whether the response was actually a sentence with inversion. However, as the task aimed at contrasts between objects/subjects there were few of such responses. Included responses were scored as correct or incorrect. Table 3 summarizes the coding scheme.

#INSERT TABLE 3

We analyzed the three subset files with complementary statistical procedures. All data were analyzed using a generalized linear mixed model with a logistic link function and

binomial variance (Baayen, 2008; Jaeger, 2008). Logistic regression is a modification of the linear model using the log odds ratio as dependent variable, instead of proportions. Regression on proportions has various mathematical drawbacks, which is why older textbooks recommend an arcsine transformation (see e.g., Ferguson & Takane, 1989). For a proportion equal to 0.5, the log odds is zero. Proportions greater than 0.5 correspond to positive log odd ratios, and proportions smaller than 0.5 to negative log odds. The transformation from log odds to proportions is non-linear: A log odds of 0 corresponds to 0.50, a log odds of 2 to 0.88, and a log odds of 4 to 0.98.

Generalized linear mixed models provide a robust statistical method for analyzing experimental data with unequal numbers of observations. Robustness of the method is relevant as the datasets included a relatively large number of missing responses due to exclusion of responses. Participant was modeled as a random-effect factor (‘Child‘), and verb form (‘VerbForm’), home language (‘HomeLanguage’), age at testing (‘Age’), L2 proficiency (‘L2Proficiency’) and clause type (‘ClauseType’) were modeled as fixed-effect factors. It is essential to model Child as a random-effect factor, as this guarantees that the model fitted to the present sample of children generalizes to the population of children from which the sample was drawn. We were unable to include item as a random-effect factor, because item-specific information was not available for all participants.

Fixed-effect factors were modeled by means of contrast between levels, where each level of a factor is contrasted to a specified reference level (treatment-coding). The levels of VerbForm - with the reference level shown in bold - were: **stem**, stem+t, stem+en. The levels of HomeLanguage were: **inflecting**, isolating. In the first subset of data we analysed, clause type was constant (main clauses without inversion). Therefore, in analyses pertaining to this subset, clause type (‘ClauseType’) was not included as a fixed-effect factor. The second

subset included two clause types: Main clauses with and without inversion. In this analysis ClauseType had the following two levels: **inversion**, no-inversion. In the third subset main clauses without inversion and relative clauses were contrasted. In analyses pertaining to this subset, the levels of ClauseType were: **main**, relative.

Age and L2 proficiency were highly correlated ($r = 0.70$). In order to decorrelate the two predictors, we predicted age by L2 proficiency, and created a new predictor ('AgeResid'), using the residuals of this model. This enabled us to determine whether any variation was explained by age, that was not explained by L2 proficiency. The correlation between age and the decorrelated age predictor was large ($r = 0.72$) indicating that the decorrelated age predictor still reflects age, albeit age to the extent that it cannot be predicted from proficiency.

We made use of a backwards elimination procedure in which predictors that did not reach significance were removed. In order to compare models, likelihood ratio tests were performed that compared the goodness of fit, while taking the costs of extra parameters into account. In this way, models were obtained that were both as simple as possible and as precise as possible.

In order to further explore the data, we applied binary recursive partitioning, buiding single classification trees and and forests (sets) of classification trees for the three datasets. Binary recursive partitioning is a non-parametric non-linear procedure that splits the data into subsets on the basis of available predictors (Hothorn, Hornik, & Zeileis, 2006). The algorithm begins with the root node, which represents the full dataset. At each node, it inspects the data for the most useful predictor, that is, the predictor that is best able to classify most observations in the subset, and splits accordingly.

One of the main differences between classification trees and linear regression models

is that the former provide more flexibility for modeling interactions (Strobl, Malley, & Tutz, 2009). Binary recursive partitioning creates one single optimized tree; such as single tree has the advantage that it provides straightforward insight into how predictors work together to explain the dependent variable. However, trees are optimized locally. Consequently, they may miss the importance of other predictors, and hence yield suboptimal predictions. Therefore, ensembles of trees are grown for random subsets of data and random subsets of predictors (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008). An ensemble of trees produces more reliable estimates of the importance of the different predictors in a given dataset. Therefore, we created single trees to explore the main trends and an ensemble of trees, a random forest, was used to evaluate the importance of the different predictors in a given dataset. A comprehensive overview of classification procedures can be found in Strobl, Malley and Tutz (2009). For an application of random forests to sociolinguistic data, the reader is referred to Tagliamonte and Baayen (2010).

Finally, in order to evaluate the goodness of fit of the mixed linear regression models, single classification trees, and ensembles of trees, we compared the predicted observations given a model/tree/ensemble of trees to the actual observations, using the index of concordance C (Chatterjee & Hadi, 2006). As a rule of thumb: A C-value above 0.80 indicates good classification performance.[5]

---

[5] All statistical analyses were conducted using R (version 2.12.0), but all methods are also available in other statistical packages. For mixed modeling LanguageR (version 1.0) was used; for classification Party (version 0.9-99991) was applied using the ctree function for single trees and the cforest function (with the control option set to unbiased) for random forests; predictive accuracy was calculated using the somers2 function from the Design (version 2.3-0) package.

Results

The impact of verb form, home language and L2 proficiency

We started with analyzing the first dataset. This dataset included all verb forms in the Dutch agreement paradigm and did not vary in word worder; all responses were SVO. Results from participants for whom no L2 proficiency score was available and who gave responses that did not fall in the categories correct or incorrect were eliminated ($N = 4$). The total number of remaining participants was 58. The number of observations comprised 685 correct responses and 158 incorrect responses. The proportions of correct responses across different conditions are in Table 4. Of the incorrect responses in 2SG, 3SG, 1PL and 3PL contexts, the proportion of omission errors was 0.9; the proportions of stem+t and stem+en substitutions were 0.08 and 0.02, respectively. The initial, full model included verb form, L2 proficiency, home language and age as predictors. Table 5 lists the estimated coefficients, their standard errors, *Z*-values and associated *p*-values for the predictors that emerged as significant in the final model.

#INSERT TABLE 4

#INSERT TABLE 5

Figures 1a-c visualize the partial effects of these predictors, back-transformed from the logits to proportions. A partial effect is the effect of a factor while all other factors are calibrated for the reference level and covariates for the median. For L2 proficiency the median was 25. In

order to illustrate the effects in an optimal manner, the reference level for verb form was set to ‘stem-t’ and for home language to ‘isolating’ in Figures 1a-c.

#INSERT FIGURES 1a-c

Figure 1a illustrates that greater proficiency offered a substantial increase in accuracy. Figure 1b shows that children with an isolating home language were less accurate than children with an inflecting home language. Figure 1c illustrates that performance with stem+t was least accurate and that performance with the bare stem was most accurate. The estimates in Table 5 and in Figure 1c suggest a difference between stem+en and stem+t. In order to determine whether or not this difference was significant, the reference level of the predictor VerbForm was set to ‘stem-t’ instead of ‘stem’; this revealed indeed a significant difference (p = 0.0014). Both contrasts survive a Bonferroni correction for multiple comparisons (both $p <$ 0.05).

Recall that the children’s proficiency highly correlated with the children’s ages. As expected, running model 1 with age as a predictor (instead of L2 proficiency), we found that a higher age predicted more correct responses ($p < 0.0001$). However, entering the decorrelated age predictor in model 1 revealed no significant effect for age ($p = 0.1405$), whereas the beneficial effect of L2 proficiency remained unaltered ($p < 0.0001$); this indicates that age did not explain variation that remained unexplained by L2 proficiency.

We explored potential interactions with the help of binary recursive partitioning, predicting the empirical log odds ratios based on correct and incorrect responses. The results are illustrated in Figure 2, which is based on one single conditional inference tree. The boxplots indicate the distributions of the log odds; the bottom and the top of the box are the

lower and upper quartiles, and the middle of the box is the median. The ends of the whiskers represent the lowest datum within the 1.5 interquartile range of the lower quartile, and the highest datum within the 1.5 interquartile range of the upper quartile. Any data not included are outliers, and plotted as dots.

#INSERT FIGURE 2

Figure 2 reveals interactions between L2 proficiency, home language and verb form. Starting at the top of the tree, which represents all data in the first subset, we see that a first split was made based on L2 proficiency: Children with a proficiency score of 28 or lower performed less accurately than children with a score above 28. The lower L2 proficiency data allowed for a second split indicating that children from isolating languages performed worse than children from richly inflecting languages. Within the data of the children with an isolating home language, a third split was made: Children with an isolating home language performed worse with stem+t than with the bare stem or stem+en. If L2 proficiency was above 28, home language and verb form were not significant factors, which was probably due to performance being at ceiling. The boxplots at the bottom of the tree represent the resulting distributions of the log odds.

Figure 3, based on a random forest analysis, provided further information on the importance of the different factors. Importance is evaluated by comparing classification accuracy for models with and without a given predictor. The greater the drop in accuracy when a predictor is not made available to the model, the more important that predictor is.The factors, or variables, are sorted: The higher its value on the x-axis, the more important a factor was in explaining variation in this dataset. L2 proficiency was the most important

factor, followed by home language and verb form. 'Child' was the least important predictor variable.

#INSERT FIGURE 3

Both the parametric logistic regression analysis and the non-parametric random forest procedure accurately predicted the observed values with *C*-scores above 0.80 (in both cases *C* = 0.90). Binary recursive partitioning was less accurate (*C* = 0.78) than the random forest procedure; this is to be expected because classification in the former case was based on only one tree whereas in the latter case many trees were grown and a voting scheme was used.

All analyses revealed that a greater L2 proficiency increased the children's accuracy with subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts in SVO sentences. Children with an inflecting home language outperformed children with isolating home languages. The effect of home language was restricted to lower L2 proficiency children. The children performed most accurately with the bare verb stem in 1SG contexts and least accurately with stem+t in 2SG and 3SG contexts. Errors were nearly all errors of omission; very few substitution errors were made. The low accuracy with stem+t was due to the lower L2 proficiency children with isolating home languages.

The role of subject-verb inversion

The second dataset included subject-verb agreement in 2SG and 3SG contexts in main clauses with (AdvVS) and without inversion (SVO). In 2SG contexts in AdvVS sentences the target form is the bare verb stem; 2SG contexts in SVO sentences target stem+t. After eliminating all datapoints for participants without correct or incorrect responses and/or

proficiency scores, data of 57 participants were analyzed, comprising 413 correct responses and 161 incorrect responses. The results of the analyses of SVO sentences revealed that the children made hardly any errors with bare verbs stems in 1SG contexts. The results in Table 6 indicate that the children performed less accurately with bare stems in 2SG contexts in AdvVS sentences:

#INSERT TABLE 6

The data in Table 6 suggest a difference between SVO and AdvVS sentences. Also, the children with an isolating home language had lower accuracies than children with inflecting home languages, except for the condition in which the bare verb stem is correct, that is, 2SG contexts in AdvVS sentences. In these sentences, the two groups had an opposite profile: Whereas the children with an isolating home language performed relatively accurately in 2SG contexts and clearly worse in 3SG contexts, the children with a richly inflecting home language performed more accurately in 3SG contexts than in 2SG contexts. In the 2SG AdvVS condition, the proportion of stem+t was 0.33 (*SD* = 0.35): Nearly all errors in this condition were thus substitutions. In the 3SG AdvVS condition, the proportion of bare verbs was 0.31 (*SD* = 0.40); here almost all errors were omissions of inflection.

Table 7 summarizes the logistic regression model fitted to the data (model 1). The initial model contained the predictors clause type, verb form, L2 proficiency, home language and age with home language and verb form entering into a significant interaction.[6]

---

[6] Note that we report the coefficients that are all simultaneously present in the model. As a consequence, the coefficients for main effects do not represent the simple main effect familiar from sequential analysis of variance tables.

#INSERT TABLE 7

Table 7 reveals that children with higher proficiencies performed more accurately; this is visualized in Figure 4a. Figure 4b sheds light on the interaction effect in Table 7: Stem+t elicited more correct responses than the bare stem for children with inflecting home languages, whereas the reverse held for children with an isolating home language.

#INSERT FIGURES 4a-b

We also considered a model that included clause type as a predictor, but this model (model 2) did not provide a better fit for the data than model 1 ($\chi 2 = 2.6591$, $df = 1$, $p = 0.1030$ (likelihood ratio tests comparing models fitted using maximum likelihood estimation)). We were unable to probe for an interaction of clause type and verb form because of empty cells. A model that included Age instead of L2 proficiency indicated that a higher age predicted more correct responses ($p = 0.0107$), but model 3, which included Age as a decorrelated predictor ($p = 0.9469$) as well as L2 proficiency ($p = 0.0001$), revealed that chronological age did not explain any variation that was left unexplained by L2 proficiency.

Binary recursive partitioning (Figure 5) indicates that home language was the primary predictor for children with an L2 proficiency score equal to or lower than 23. In this lower L2 proficiency group, children with an isolating home language performed worse than children with an inflecting home language. For children with higher L2 proficiencies (above 23) verb form was the primary predictor: Performance was better with stem+t than with the bare stem form. Performance with stem+t led to another split based on home language. Children with a

higher L2 proficiency who spoke an isolating home language performed worse with stem+t than children in the same L2 proficiency group who spoke an inflecting language at home. The boxplots at the bottom illustrate the distribution of the log odds.

#INSERT FIGURE 5

Figure 6 illustrates the importance of the factors according to a random forest analysis. Home language emerged as the most important predictor followed by L2 proficiency, verb form and, at a distance, the effect of individual children. These results are in line with the logistic regression model. Logistic regression ($C = 0.86$) and random forest classification ($C = 0.85$) predicted the observed values accurately; the best single binary recursive partitioning tree was less accurate ($C = 0.76$), as expected.

#INSERT FIGURE 6

Analyses of performance with subject-verb agreement in 2SG and 3SG contexts in SVO and AdvVS sentences indicated that children with higher L2 proficiency made fewer errors than lower L2 proficiency children. Children with isolating home languages had lower accuracies than children with inflecting home languages, in particular in lower L2 proficiency groups. Effects of verb form were clearer in higher L2 proficiency children. Children with inflecting home languages used stem+t forms instead of the bare verb stem in 2SG contexts. Children with isolating home languages tended to drop the suffix *-t* in obligatory contexts.

Sentence position

The last dataset comprised data on children's accuracy with 3SG contexts in SVO and SOV sentences. In this dataset, there were 448 correct responses and 106 incorrect responses (*N* participants = 57). Of the incorrect responses, the proportions of omissions and substitutions were 0.94 and 0.06, respectively. In both word order conditions, the same verb form was targeted (stem+t), and the only difference was clause type, which involved differences in the position of the verb in the sentence (final versus medial position). Table 8 lists the accuracies across conditions:

#INSERT TABLE 8

As before, we started with the full model that included, apart from clause type, L2 proficiency, home language and age as predictors. Table 9 presents the coefficients of the best fitting logistic model, regressing the correctness of 3SG. The model estimates in Table 9 indicate that children with an inflecting home language performed better than children with an isolating home language. Children with higher L2 proficiencies performed better with subject-verb agreement in relative clauses than in main clauses, whereas the reverse held for lower proficiency children ($p = 0.0431$). Figures 7a-b visualize the partial effects of the model. The model with age instead of L2 proficiency did not reveal a significant main effect for age ($p = 0.0536$) or an interaction between age and clause type ($p = 0.0921$).

#INSERT TABLE 9

#INSERT FIGURES 7a-b

Figure 8 shows the outcome of the binary recursive partitioning procedure by indicating the distribution of the log odds. This method confirms that more proficient children performed significantly better with 3SG agreement in SOV than in SVO sentences. An interaction of L2 proficiency and home language emerged as well: For those participants with L2 proficiency scores of 23 or below, participants who speak an inflecting home language outperformed participants with an isolating home language.

#INSERT FIGURE 8

Figure 9 illustrates the relative importance of the variables. L2 proficiency emerged as the most important predictor, followed by home language, clause type and the effect of individual children. All three procedures predicted the observed values accurately: For logistic regression a *C* of 0.96 was obtained, whereas for binary recursive partitioning $C = 0.84$ and for random forest $C = 0.90$.

#INSERT FIGURE 9

More proficient children made fewer errors with stem+t in 3SG contexts than less proficient children. Children with isolating home languages performed worse than children with inflecting home languages, especially if they had lower L2 proficiency. Higher L2 proficiency children more often omitted inflection in SVO sentences than in SOV sentences.

Discussion

We presented the results of a study in which we tested L2 Dutch children's accuracy with subject-verb agreement. Participants were assigned to two groups, one with children who spoke at home an isolating language (Mandarin, Cantonese), the other with children whose home language had rich inflection (Moroccan-Arabic/Tarifit, Turkish). The first dataset included all three verb forms of the Dutch present tense paradigm (stem, stem+t, stem+en) and was restricted to main clauses without inversion (SVO). Accuracy with subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts was tested. The second dataset compared 2SG and 3SG contexts across main clauses without inversion (SVO) and main clauses with inversion (AdvVS). The third dataset dealt with 3SG subject-verb agreement in main clauses without inversion (SVO) and relative clauses (SOV). All analyses consistently showed that children who were more proficient in the L2 performed better with subject-verb agreement than less proficient children. Unsurprisingly, the L2 proficiency scores were highly correlated with the children's ages at time of testing, but in this dataset L2 proficiency was the more important predictor. We also observed that home language, verb form and sentence position impacted the children's performance.

Verb form

Our first research question was: To what extent does verb form - bare or inflected - impact L2 children's accuracy with subject-verb agreement? The effects of verb form are relevant for assessing the MISH. According to this hypothesis, the errors with subject-verb agreement made by L2 children reflect application of a default rule, which results in the incorrect use of bare verbs (Haznedar, 2001; Haznedar & Schwartz, 1997; Ionin & Wexler, 2002; Prévost, 2003). The MSIH thus predicts that children learning L2 Dutch will make

errors in conditions that target stem+t or stem+en, but not in the conditions that target bare verbs, and that these errors comprise omissions of inflection.

This prediction is largely confirmed by the analyses of the first dataset. Analyses of dataset 1 indicated that children in both L1 groups made hardly any errors in 1SG contexts, where the bare verb is correct. Significantly more errors were made in those contexts that involved inflected verb forms. In these contexts (2SG, 3SG, 1PL, 3PL) omission errors predominated, and very few substitution errors occurred. The outcomes of the second dataset indicate that in more exceptional contexts where the bare verb is grammatical, such as 2SG contexts in sentences with inversion, performance with the bare stem was relatively poor, and substitutions were frequent. This observation seems more difficult to reconcile with the MSIH.

At this point, some more background information on Dutch morphology becomes relevant. Given the premises of the MSIH, the following question arises: Why is the bare stem allowed, even obligatory, in 2SG contexts in Dutch sentences with inversion? The MSIH adopts the premises of Distributed Morphology, a morphological theory in which obedience to a blocking principle leads to use of the most specific paradigmatic form given contextual restrictions (in this case, restricting properties of the subject). Thus, in 2SG contexts, a morphological form associated with 2SG features is used (stem+t). According to Distributed Morphology, the unexpected, but systematic, occurrence of the default form can happen due to language-specific, context-sensitive rules that “impoverish” contextual information (Bonet, 1991; Halle, 1997). Without going into any detail, such a rule removes one feature in the presence of another, creating a licit context for the default rule. In order to account for bare stem forms in 2SG inversion contexts, it has been argued that Dutch has an impoverishment rule that removes person features in 2SG contexts if verb and 2SG subject

are in the same prosodic phrase, which happens in inversion sentences (Ackema & Neeleman, 2004). If this rule applies, the stem+t form is too specific and for that reason ruled out.

The substitution errors in 2SG contexts could be made by L2 children who did not know the impoverishment rule, and, as a result, used the specific form (stem+t) instead of the default form (stem). Three participants consistently produced incorrect stem+t in 2SG inversion contexts; these children may have been unaware of the impoverishment rule. 17 participants showed variable behaviour in 2SG contexts in sentences with inversion, sometimes using the bare stem and on other occassions using stem+t. It is possible that these children did not know the impoverishment rule but dropped inflection due to performance factors, similar to 2SG contexts in main clauses without inversion. However, within the subset of 17 participants with variable behaviour, (correct) bare verbs in 2SG sentences with inversion (50%) were more frequent than (incorrect) bare verbs in 2SG sentences without inversion (20%). This difference suggests that variability in 2SG inversion sentences is caused by different factors than variability in 2SG sentences without inversion. Thus, also at closer inspection, it is not immediately clear how the error pattern in 2SG contexts in sentences with inversion could be explained under the MSIH.

One observation that may provide more insight into the children's error patterns is that L1 isolating children tend to drop inflection more frequently in 3SG contexts in sentences with inversion than in 3SG contexts in sentences without inversion (Table 6). This suggests that being exposed to bare verbs in both 1SG and 2SG contexts in inversion sentences in the input spreads to 3SG contexts in the speech of L1 isolating children. This effect is reminiscent of L2 English children with isolating L1s who drop 3SG *-s* more frequently with verbs that appear often in bare form in the input (Blom et al., in press).

The more general observation underlying both observations would be that L1 isolating children are more sensitive to bare verb forms in the input than L1 inflecting children, suggesting the emergence of a learning bias from prior learning (Ellis, 2006). Prior to learning the L2, children with an isolating L1 have learned a language with only bare verbs. The resulting attention to bare verbs in the L2 may explain two patterns in the data. First, for L1 isolating children hearing the bare verb form in certain contexts (e.g., 1SG and 2SG inversion) may result in overuse of the bare verb in overlapping contexts (e.g., 3SG inversion). Secondly, it could lead to a relatively strong association between 2SG subjects and the bare verb in inversion contexts for this group of children, and hence to relatively high accuracies in this rather exceptional context.

Following up on this intuition, we briefly explore the hypothesis that substitution errors in 2SG inversion sentences are influenced by the frequency of co-occurrence of subject form and verb form. Evidence for such input effects has been reported for subject-verb agreement in L1 English children (Joseph, Serratrice, & Conti-Ramsden, 2002; Wilson, 2003). Suppose that L2 children, like L1 children, are sensitive to the overlap between different structures in the input (Tomasello, 2003), and that L2 children make use of distributional contingencies - the likelihood of x given y - in the input, similar to what has been proposed for the L1 acquisition of grammatical dependencies between words (Newport & Aslin, 2004; Santelmann & Jusczyk, 1998).

The expectation would be that in the Dutch input the contingency between stem+t and 2SG subjects will be stronger than the contingency between bare verbs and 2SG subjects, because in spoken Dutch, main clauses without inversion outnumber main clauses with inversion at a ratio of 2.3:1 (based on analyses performed on the *Corpus Spoken Dutch* reported in Van der Wouden, Schuurman, Schouppe, & Hoekstra, 2003). Also, the association

between 1SG subjects and bare verbs will be stronger than the association between 2SG subjects and bare verbs, because 1SG subjects always appear with bare verbs, whereas 2SG subjects will also often co-occur with stem+t forms in overlapping constructions (with the same verb, for instance).

At first sight, this hypothesis makes the correct predictions that more errors occur in 2SG contexts in inversion sentences than in 2SG contexts in sentences without inversion and that L2 children perform better with bare verbs in 1SG contexts than in 2SG contexts in inversion sentences. However, more research is needed to evaluate the role of frequency distributions and its interplay with L1 properties. Such a study should be based on frequency counts of the different verb forms across contexts in a corpus that is representative of the children's input.

Sentence position

The role of sentence position pertains to our second research question: Does sentence position - medial versus final - influence L2 children's accuracy with subject-verb agreement? We found support for an effect of sentence position: Omission rates of 3SG *-t* were higher in medial position (SVO) than in sentence final position (SOV) for children with greater L2 proficiency. Song et al. (2009) found similar effects of position for 3SG *-s* in L1 English. They suggested that medial position created fewer opportunities to produce the agreement morpheme and posed articulatory challenges for the child. They also argued that lengthening of 3SG *-s* in final position increased input salience.

In our study, lengthening did not play a role because the target form ended with a stop. However, differences in time to (prepare to) produce [t] in medial and in final position could have led to relatively more drop in medial position. This effect may have become more

pronounced through phonological complexity and articulatory challenges that resulted from -*t* suffixation. None of the verbs in our study ended on a vowel, reflecting the general avoidance in Dutch of vowel-final verbs (Trommelen, 1989). Therefore, suffixation of *–t* created complex coda clusters. Such clusters are more prone to drop than singleton coda's, in particular under time restrictions (Song et al., 2009). In addition to coda complexity, difficulties with articulatory planning may have contributed to drop of *-t* when words followed the verb, as was the case with verbs in medial position.

Properties of relative clauses other than word order may have had some influence as well. In the SOV condition, the children's task was to complete a sentence that started with, for instance, *Dit is de man die* … ('This is the man who …'). Such a structure, in which the relative clause narrows down the possible interpretations of the sentence subject, is typically used with third person subjects. Relative clauses may thus create a facilitating context for third person verb forms. Moreover, there may have been a phonological effect: The final [t] in the demonstrative subjects *dit* and *dat* could have triggered use of stem+t. Evidence for phonological priming effects has been reported for English irregular past tense (Stemberger, 2004).

Home language, L2 proficiency and age

The last two research questions of our study address individual differences: Is the development of subject-verb agreement determined by inflectional properties of L2 children's home language? To what extent is the development of subject-verb agreement influenced by L2 children's age and their proficiency in the L2? A clear effect of L2 children's home language was found, as well as an interaction of home language and L2 proficiency. Children with a richly inflecting home language (Moroccan-Arabic/Tarifit or Turkish) consistently

outperformed children with an isolating home language (Cantonese or Mandarin). Binary recursive partitioning revealed that home language impacted children with lower L2 proficiency. This outcome is straightforward: The more advanced the children are in the L2, the less influence their home language will have.

The impact of home language in our study is in line with recent studies on the acquisition of verb inflection in child L2 English (Blom et al., in press; Paradis, in press). These effects suggest that L2 children transfer properties related to inflection from their L1 to their L2, in line with the Full Access Full Transfer hypothesis proposed for L2 adults (Schwartz & Sprouse, 1996; White, 2003). The Full Access Full Transfer hypothesis is phrased in structural terms (e.g., parameter-settings), but the effect of home language may also be related to perception or attention. For instance, children who are familiar with a pro-drop language in which verbal inflection provides cues for identifying the sentence subject (such as Turkish, Moroccan-Arabic/Tarifit but not Mandarin and Cantonese) will be attentive to verbal inflection in the L1 and they may transfer this habit to their L2. Or, as pointed out above, L2 children may be more sensitive to the distribution of a particular form, dependent on inflectional properties of the L1.

L2 Dutch children from isolating languages experienced particular difficulties with stem+t. As explained earlier, suffixation of *-t* resulted in complex coda clusters. Given that both Cantonese and Mandarin do not allow word-final consonant clusters - Mandarin does not even permit word-final [t] -, the children's difficulties with stem+t could reflect effects of transfer of L1 phonotactic restrictions in the earlier stages of development. In this respect, it is relevant to note that children with an isolating home language performed worse with stem +t than with stem+en, which is a syllabic morpheme. Unlike suffixation of *-t,* suffixation of *-en* will not result in a word-final consonant cluster. Support for this analysis comes from

studies that report effects of phonological transfer in bilingual children (Flege, Munro, & MacKay, 1995; Fabiano-Smith & Barlow, 2009; Fabiano-Smith & Goldstein, 2005), and in particular from studies which show that bilingual children's production of consonant clusters in onset position is affected by transfer (Yavas & Barlow, 2006).[7]

Two previous studies that addressed effects of transfer on subject-verb agreement in child L2 production investigated the production of *-s* and *-ed* in child L2 English, but did not find any L1 differences (Paradis, 2005; Paradis et al., 2008). Possibly, effects of transfer are highest at an intermediate stage of proficiency, when the developmental readiness for transfer is present, but the learners are not so proficient yet that their performance would be at ceiling and not reveal any influence of the L1 anymore (Wode, 1983; Zobl, 1980). The lower proficiency L2 Dutch children in our study may have been in this stage, in contrast to the L2 English children in Paradis' study. In the latter study the children's length of exposure to the L2 was relatively short (mean exposure in months = 9.5, range of exposure in months = 2-18), which may support this idea. Also, collapsing *-s* and *-ed* may have decreased the impact of L1. Children who have an isolating L1 might be more aware of past tense marking than of 3SG *-s* in English, given that past tense conveys semantic information whereas 3SG *-s* in English is an agreement marker.

L2 proficiency was a more important factor than age, and age did not explain variation that was unexplained by L2 proficiency. The older children attended school longer

---

[7] One reviewer correctly pointed out that this hypothesis predicts that the L1 isolating children would show general difficulties with complex coda clusters and that nouns with complex coda clusters, for instance, will also exhibit cluster reduction strategies. Our experimental set-up, which was developed to elicit finite verbs, did not allow testing this prediction.

than younger children. Hence, they were not only cognitively more mature than younger children, but they were also more experienced users of the L2. L2 proficiency may index both L2 experience and cognitive maturity, but the greater importance of L2 proficiency (in comparison to age) points to other factors as well. Most likely, L2 proficiency also contains information about language aptitude. Language aptitude is a reliable predictor of success in adult L2 development (Dörnyei & Skehan, 2003), and the same may hold for child L2 development. Language aptitute will influence both the children's L1 and L2 development, and hence, it would be expected that L1 proficiency predicts L2 proficiency (Castilla, Restrepo, & Perez-Leroux, 2009). In the present study, no information about L2 children's L1 skills was collected. It could be insightful, however, to relate effects of transfer not only to measures of L2 development (e.g., length of exposure to the L2, L2 proficiency), but also to L1 development and language dominance profiles. Such a method could shed further light on the different factors that condition effects of transfer in child L2 acquisition.

Applied issues

The omission of 3SG *-s* is a prototypical error made by English-learning children with a specific language impairment (cf. Bedore & Leonard, 1998; Rice & Wexler, 1996). Dutch children with specific language impairment who have the same age as the children in our study overuse bare verbs stems as well, and continue doing so at later ages (De Jong, 1999; Steenge, 2006). Steenge (2006, p. 54) even claims that incorrect bare verbs are a clinical marker for specific language impairment in Dutch, regardless of whether Dutch is a child's L1 or L2. However, the data presented here show that non-impaired Dutch L2 children make errors in contexts that require overt inflection, in particular if their L1 is isolating. This was particularly clear in 2SG and 3SG contexts in main clauses without inversion, and in 3SG

contexts in sentences with inversion. Therefore, if agreement errors are used as a clinical marker, L1 background is a factor that definitely should be taken into account. In this respect, we echo the conclusion of Jia and Fuse (2007, p. 1297): The results call for caution when interpreting and using morphological errors as indications of speech, language or learning disorders in the immigrant children population merely on the basis of similarities between the error profiles of non-impaired L2 children and children diagnosed with specific language impairment.

Conclusion

The aim of this study was to investigate factors modulating Dutch L2 children's use of subject-verb agreement. Analyses of results of controlled elicitation experiments revealed effects of verb form, sentence position, home language and L2 proficiency. Most errors were omissions of agreement inflection, in line with the Missing Surface Inflection Hypothesis. However, in more exceptional contexts, substitution errors were made as well, which was more difficult to reconcile with this hypothesis. Inflection drop was more frequent in medial than in final sentence position, which was interpreted as evidence for the impact of time restrictions in combination with phonological complexity and articulatory planning on L2 children's expression of subject-verb agreement. Effects of transfer emerged in children with lower L2 proficiency: Children with richly inflecting home languages were more accurate than children with isolating home languages. Transfer of phonotactic constraints on coda complexity could explain specific between-group differences. Finally, L2 proficiency was highly correlated with age, but age did not predict variation that was unexplained by L2 proficiency.

Appendix: Examples of picture description task and activity description task

Picture description task

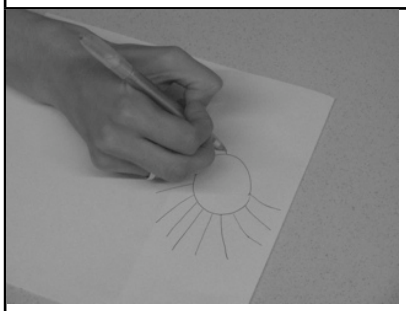
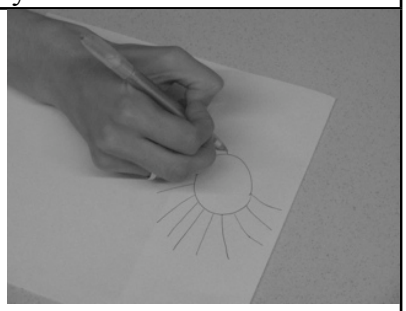
3SG condition (main clauses no inversion, relative clauses)

3PL condition (main clauses no inversion)

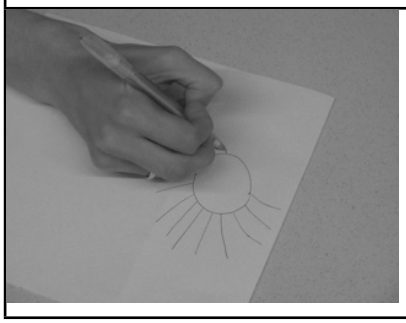
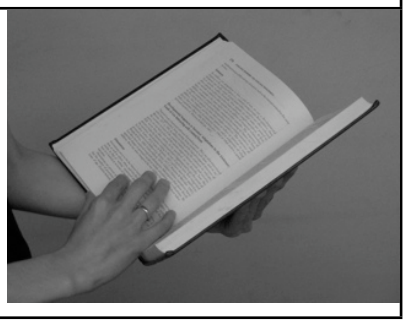
Appendix: Examples of picture description task and activity description task

Activity description task

1PL condition: Experimenter and participant pick up the same card and act out the same activity.

1SG/2SG condition: Experimenter and participant pick up a different card and act out different activities.

References

Ackema, P. & Neeleman, A. (2004). Beyond morphology. Interface conditions on word formation. Oxford: Oxford University Press.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bailey, N., Madden, C. & Krashen, S. (1974). Is there a 'natural sequence' in adult second language learning? *Language Learning*, *24*, 235-243.

Bedore, L. & Leonard, L. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of speech, language and hearing research*, *41*, 1185-1192.

Blom, E. & De Korte, S. (2011). Dummy auxiliaries in child and adult second language acquisition of Dutch. *Lingua*, *121*, 906–919

Blom, E., Orgassa, A. & Polišenská, D. (2008). FlexiT: Elicitatie-materiaal voor de inventarisatie van de mondelinge taalvaardigheid van werkwoordsvervoeging, werkwoordsplaatsing, vervoeging van het bijvoeglijk naamwoord en lidwoordtoekenning bij diverse taalleerders van het Nederlands [Elicitation material for the collection of oral profiency data on verbal inflection, verb placement, adjectival inflection and definite articles in various learners of Dutch]. University of Amsterdam.

Blom, E ., Polišenská, D. & Weerman, F. (2008). Articles, adjectives and age of onset: the acquisition of Dutch grammatical gender. *Second Language Research*, *24(3)*, 289-323.

Blom, E., Paradis, J. & Sorenson-Duncan, T., in press. Effects of input properties, vocabulary size and L1 on the development of third person singular –s in child L2 English. *Language Learning*.

Blumenthal, M. & Julien, M. (2000). Diagnostiek van spraak- en taalproblemen bij meertalige kinderen, Geen diagnose zonder anamnese meertaligheid. *Logopedie en Foniatrie*, 13-17.

Bonet, E. (1991). Morphology after syntax: Pronominal clitics in Romance, Doctoral dissertation MIT, Cambridge MA (distributed by MIT Working Papers in Linguistics).

Carroll, J. B. & Sapon, S. (2002). Modern Language Aptitude Test: Manual 2002 Edition. Bethesda, MD: Second Language Testing, Inc.

Castilla, A. P., Restrepo, M. A. & Perez-Leroux, A. T. (2009). Individual differences and language interdependence: A study of sequential bilingual development in Spanish-English preschool children. *International Journal of Bilingual Education and Bilingualism*, *12(5)*, 565-580.

Chatterjee, C. & Hadi, A. S. (2006). *Regression Analysis by Example (4rth edition)*. New York: Wiley.

Chilla, S. (2008). Erstsprache, Zweitsprache, Spezifische Sprachentwicklungsstörung? Eine Untersuchung des Erwerbs der deutschen Hauptsatzstruktur durch sukzessiv-bilinguale Kinder mit türkischer Erstsprache [L1, L2, SLI? The acquisition of German sentence structure by successive-bilingual Turkish-German children]. Hamburg: Dr. Kovac.

Dalal, R. H., & Loeb, D. F. (2005). Imitative production of regular past tense –ed by English-speaking children with specific language impairment. *International Journal of Language and Communication Disorders*, *40*, 67–82.

De Jong, J. (1999). Specific language impairment in Dutch: inflectional morphology and argument structure. Unpublished doctoral dissertation, Groningen University.

Dörnyei, Z. & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty & M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 589–630). Oxford: Blackwell.

Dulay, H. & Burt, M. (1973). Should we teach children syntax? *Language Learning*, *23*, 245-258.

Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, *24*, 37-53.

Eisenbeiss, S. (2010). Using production methods in language acquisition research. In E. Blom & S. Unsworth (Eds), *Experimental methods in language acquistion research* (pp. 11-34). Amsterdam/Philadelphia: Benjamins.

Ellis, N.C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking and perceptual learning. *Applied Linguistics*, *27(2)*, 164-194.

Fabiano-Smith, L. & Barlow, J.A., (2009). Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism*, 1-17.

Fabiano-Smith, L. & Goldstein, B. (2005). Phonological cross-linguistic effects in bilingual Spanish-English speaking children. Journa*l of Multilingual Communication Disorders*, *3(1)*, 56-63.

Ferguson, G.A. & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th Edition). New York [etc.]: McGraw-Hill Book Company.

Flege, J.E., Munro, M.J., & MacKay, I.R.A . (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, *97*, 3125-3134.

Goad, H. & White, L. (2004). Ultimate attainment of L2 Inflection. Effects of L1 prosodic structure. In Foster-Cohen, M. Sharwoord-Smith, A. Sorace & M. Ota (Eds.) *EUROSLA Yearbook 4* (pp. 119-145). John Benjamins: Amsterdam/Philadelphia,.

Goad, H. & White, L. (2008). Prosodic structure and the representation of L2 functional morphology: A nativist approach. *Lingua*, *118*, 577-594.

Goldschneider, J. & DeKeyser, R. (2001). Explaining the "Natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, *51(1)*, 1-50.

Halle, M. (1997). Distributed Morphology: Impoverishment and Fission, *MIT Working Papers in Linguistics*, *30*, 425-449.

Halle, M. & Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S. J. Keyser (Eds), *The View from Building 20, Essays in Linguistics in honour of Sylvain Bromberger* (pp. 111-176). MIT Press: Cambridge MA.

Haznedar, B. (2001). The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition*, *23*, 1-39.

Haznedar, B. & Schwartz, B. (1997). Are there Optional Infinitives in child L2 acquisition? In E. Hughes, M. Hughes and A. Greenhill (Eds), *Proceedings of the 21st annual Boston University Conference on Language Development* (pp. 257-268). Somerville MA: Cascadilla Press.

Herschensohn, J., Stevenson, J. & Waltmunson, J. (2005). Children's acquisition of L2 Spanish morphosyntax in an immersion setting. *International Review of Applied Linguistics*, *43*, 193-217.

Hoogland, J. (1996). *Marokkaans Arabisch, een cursus voor zelfstudie en klassikaal gebruik*. Amsterdam: Bulaaq.

Hothorn, T, Hornik, K. & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, *15(3)*, 651-674.

Ionin, T. & Wexler, K. (2002). Why is 'is' easier than 's'? Acquisition of tense/agreement morphology by child L2-English learners. *Second Language Research*, *18(2)*, 95-136.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59(4)*, 434-446.

Jia, G. & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: age-related differences. *Journal of speech, language and hearing research*, *50*, 1280-1299.

Joseph, K., Serratrice, L., & Conti-Ramsden, G. (2002). Development of copula and auxiliary BE in children with specific language impairment and younger unaffected controls. *First Language*, *22*, 137-172.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, *25(2)*, 173-227.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, *9*, 409-420.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, *26*, 125-134.

Lewis, G. L. (1967). *Turkish grammar*. Oxford: Clarendon.

Lin, H. (2001). *A grammar of Mandarin Chinese*. Munich: Lincom Europa.

Matthews, S. & Yip, V. (1991). *Cantonese: A comprehensive grammar*. London: Routledge.

McDonald, J. (2000). Grammaticality judgments in a second language: influences of age of acquisition and native language. *Applied Psycholinguistics*, *21*, 395-423.

Meisel, J.M. (2009). Second Language Acquisition in Early Childhood. *Zeitschrift für Sprachwissenschaft*, *28*, 5-34.

Newport, E.L. & Aslin, R.N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127-162.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, speech and hearing services in the schools*, *36*, 172-187.

Paradis, J. (in press). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic approaches to bilingualism*.

Paradis, J., Rice, M., Crago, M. & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, *29*, 689-722.

Prévost, P. (2003). Truncation and missing surface inflection in initial L2 German. *Studies in Second Language Acquisition*, *25*, 65-97.

Rice, M. & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of speech, language, and hearing Research, 39*, 1239-1257.

Santelmann, L. & Jusczyk, P. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for processing limitations. *Cognition*, *69*, 105-134.

Schwartz, B.D. (2004). Why Child L2 acquisition?. In: J. van Kampen & S. Baauw (Eds.), *The proceedings of GALA 2003* (pp. 47-66). Utrecht: LOT Occasional Series.

Schwartz, B. D. & Sprouse, R.A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, *12*, 40-72.

Song, J.Y, Sundara M. & Demuth, K. (2009). Phonological constraints on children's production of English third person singular –s. *Journal of speech, language and hearing research*, *52*, 623-642.

Steenge, J. (2006). Bilingual children with specific language impairment: Additionally disadvantaged? Unpublished doctoral dissertation, Radboud University Nijmegen.

Stemberger, J. P. (2004). Phonological priming and irregular past. *Journal of Memory and Language*, *50(1)*, 82-95.

Strobl, C., Mallay, J. & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forrests. *Psychological Methods*, *14(4)*, 323-348.

Strobl, C., Boulesteix, A-L, Kneib, T, Augustin, T. & Zeileis, A. (2008). Conditional variable importance for random forests. BMC Bioinformatics, 9, 307. http:// www.biomedcentral.com/1471-2105/9/307.

Tagliamonte, S. & Baayen, R. H. (2010). Models, forests and trees of York English: Was/were variation as a case study for statistical practice. Manuscript submitted for publication.

Tomasello, M. (2003). *Constructing a language*. Cambridge MA: Harvard University Press.

Trommelen, M. (1989). Lettergreepstruktuur en woordkategorie. *De Nieuwe Taalgids*, *82*, 64-77.

Unsworth, S. & Blom, E. (2010). Comparing L1 children, L2 children and L2 adults. In E. Blom & S. Unsworth (Eds). *Experimental methods in language acquistion research* (pp. 201-222). Amsterdam/Philadelphia: Benjamins.

Vainikka, A. & Young-Scholten, M. (2010). All acquisition begins with the projection of a bare verb phrase. *Applied Psycholinguistics*, *31(2)*, 332-339.

Van der Wouden, T., Schuurman, I., Schouppe, M. & Hoekstra, H. (2003). Harvesting Dutch trees: Syntactic properties of spoken Dutch. In T. Gaustad (Ed.), *Computational linguistics in the Netherlands 2002: selected papers from the thirteenth CLIN meeting* (pp. 129-142). Amsterdam/New York: Rodopi.

Verhoeven, L. & Vermeer, A. (2002). *Taaltoets Alle Kinderen* [Children's language assessment]. Tilburg: Zwijsen.

Wexler, K., Schaeffer, J. & Bol, G. (2004). Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play an important role in morphological theory. *Syntax*, *7(2)*, 148-198.

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge University Press, Cambridge.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of child language*, *30*, 75-115.

Wode, H. (1983). One the systematicity of L1 transfer in L2 acquisition. In H. Wode (Ed.), *Papers on language acquisition, language learning and language teaching* (pp. 144-149). Heidelberg: J. Groos,.

Yavaş, M. & Barlow, J. (2006). Acquisition of #sC clusters in Spanish-English bilingual children. *Journal of Multilingual Communication Disorders*, *4(3)*, 182 – 193

Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research*, *24(2)*, 227-250.

Zobl, H. (1980). The formal and developmental selectivity of L1 influence in L2 acquisition. *Language learning*, *30*, 43-57.

Table 1: Datasets

Relative clause (SOV)	Main clause without inversion (SVO)	Main clause with inversion (AdvVS)
	1SG (stem)	
	2SG (stem+t)	2SG (stem)
3SG (stem+t)	3SG (stem+t)	3SG (stem+t)
	1PL (stem+en)	
	3PL (stem+en)	

Table 2: Participant information

Country of descent	China	Morocco	Turkey
Home language	Mandarin/Cantonese	Moroccan-Arabic/Tarifit	Turkish
Number	20	15	27
Boys	8	9	14
Age at testing *M*	6;5	6;0	6;5
Age at testing *SD*	1.54	1.48	0.95
Age at testing range	4;04 – 9;11	4;02 – 8;04	4;08 – 8;05
L2 Proficiency *M*	24.45	23	26.88
L2 Proficiency *SD*	8.42	8.81	7.19
L2 Proficiency range	9 – 37	5 – 36	9 – 38

Table 3: Coding of correct and incorrect responses

Condition	Correct verb form	Incorrect verb form
1SG SVO	stem	stem+t, stem+en
2SG SVO	stem+t	stem, stem+en
2SG AdvVS	stem	stem+t, stem+en
3SG SVO, SOV, AdvVS	stem+t	stem, stem+en
1PL SVO	stem+en	stem, stem+t
3PL SVO	stem+en	stem, stem+t

Table 4: Mean proportions correct (*SD*) in first dataset

Condition	Verb form	$M_{\text{prop.correct}}$ Overall	$M_{\text{prop.correct}}$ Isolating	$M_{\text{prop.correct}}$ Inflecting
1SG	stem	0.97 (0.16)	0.94 (0.25)	0.98 (0.09)
2SG	stem+t	0.77 (0.33)	0.58 (0.43)	0.86 (0.29)
3SG	stem+t	0.74 (0.32)	0.54 (0.36)	0.85 (0.23)
1PL	stem+en	0.85 (0.30)	0.83 (0.25)	0.86 (0.34)
3PL	stem+en	0.81 (0.30)	0.70 (0.36)	0.88 (0.25)

Table 5: Model 1 estimates, standard errors, z-values and associated p-values; accuracy of subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts in main clauses without inversion (SVO)

	Estimate	*SE*	*Z*	*P*
(Intercept)	2.72	1.01	2.7	0.0069
L2 Proficiency	0.11	0.02	4.27	0.0000
Home language (isolating)	-1.37	0.39	-3.53	0.0004
Verb form (stem-en)	-2.72	0.78	-3.49	0.0005
Verb form (stem-t)	-3.42	0.78	-4.39	0.0000

Table 6: Mean proportion correct (*SD*) in second dataset

Condition	Verb form	$M_{\text{prop.correct}}$ Overall	$M_{\text{prop.correct}}$ Isolating	$M_{\text{prop.correct}}$ Inflecting
2SG - AdvVS	stem	0.64 (0.36)	0.83 (0.32)	0.58 (0.36)
3SG - AdvVS	stem+t	0.66 (0.40)	0.37 (0.39)	0.80 (0.32)
2SG - SVO	stem+t	0.77 (0.33)	0.58 (0.43)	0.86 (0.29)
3SG - SVO	stem+t	0.74 (0.32)	0.54 (0.36)	0.85 (0.23)

Table 7: Model 1 estimates, standard errors, z-values and associated p-values; accuracy with subject-verb agreement in 2SG and 3SG contexts in main clauses with inversion (AdvVS) and without inversion (SVO); the asterisk indicates the interaction of two predictor variables

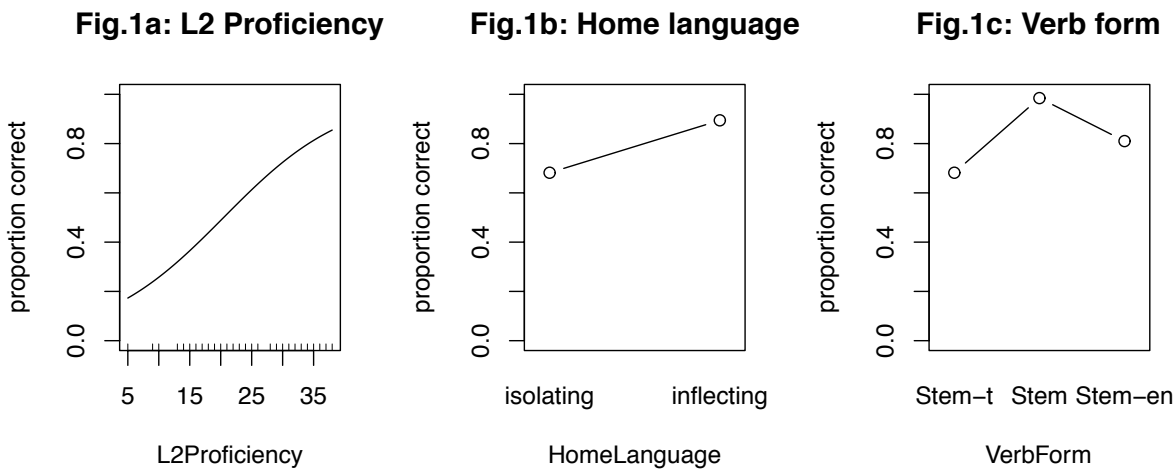
	Estimate	*SE*	*Z*	*P*
(Intercept)	-1.53	0.56	-2.73	0.0064
L2 Proficiency	0.07	0.02	3.88	0.0001
Home language (isolating)	1.30	0.68	1.91	0.0567
Verb form (stem-t)	1.51	0.32	4.70	0.0000
Verb form (stem-t) * Home language (isolating)	-2.94	0.70	-4.23	0.0000

Table 8: Mean proportion correct (*SD*) in third dataset

Condition	Verb form	$M_{\text{prop.correct}}$ Overall	$M_{\text{prop.correct}}$ Isolating	$M_{\text{prop.correct}}$ Inflecting
3SG SVO	stem+t	0.74 (0.32)	0.54 (0.36)	0.85 (0.23)
3SG SOV	stem+t	0.87 (0.25)	0.78 (0.31)	0.90 (0.23)

Table 9: Model 1 estimates, standard errors, z-values and associated p-values; accuracy with subject-verb agreement in 3SG contexts in main clauses without inversion (SVO) and relative clauses (SOV); the asterisk indicates the interaction of two predictor variables

	Estimate	*SE*	*Z*	*P*
(Intercept)	0.20	0.79	0.26	0.7949
L2 Proficiency	0.07	0.03	2.43	0.0151
Home language (isolating)	-1.43	0.43	-3.34	0.0008
Clause type (relative)	-0.62	0.85	-0.73	0.4673
L2 Proficiency * Clause type (relative)	0.07	0.03	2.02	0.0431



Figures 1a-c: Partial effects model 1 (mixed logistic regression); accuracy of subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts in main clauses without inversion (SVO)

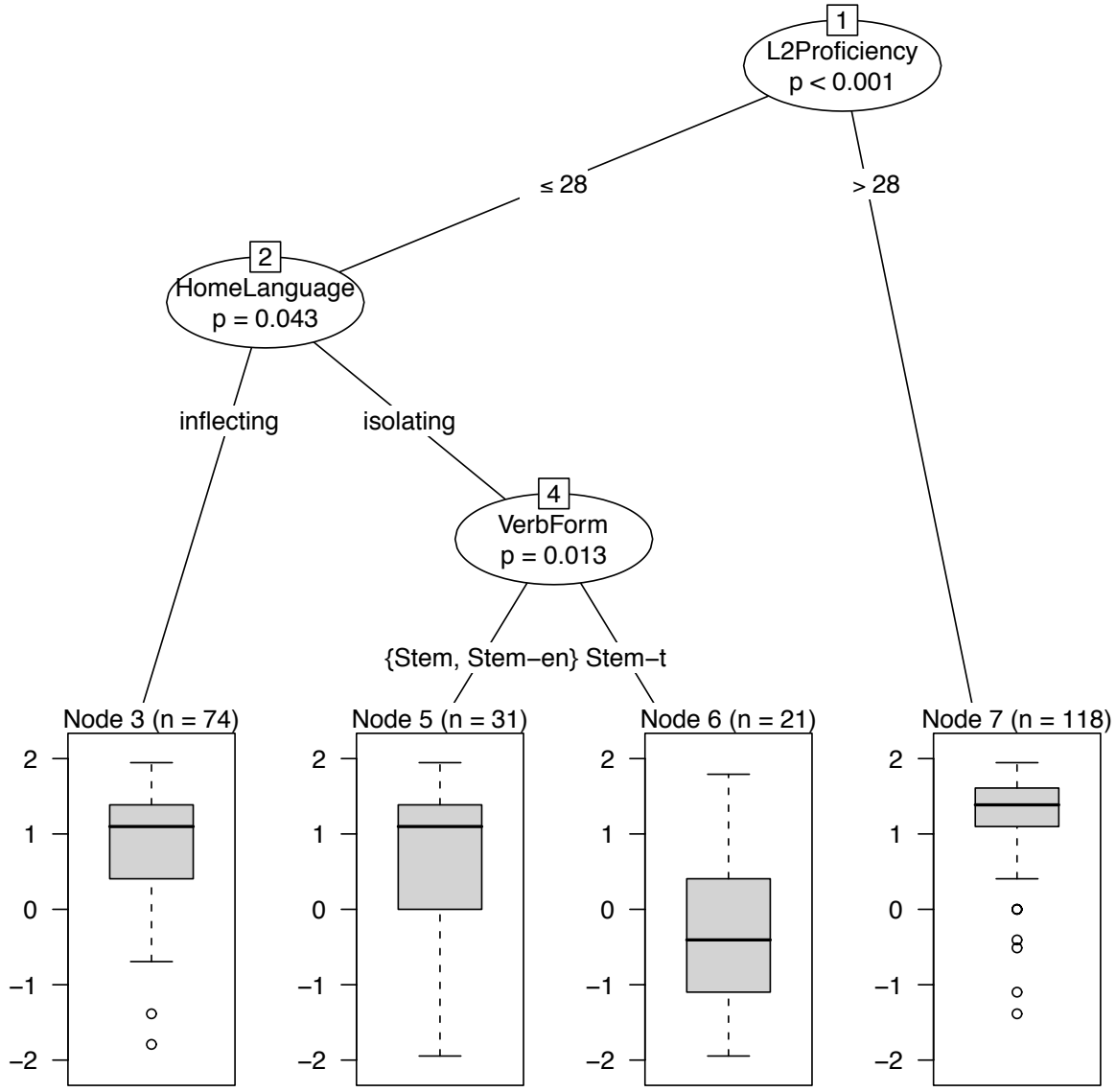


Figure 2: Conditional inference tree, predicting the empirical log-odds ratios; accuracy of subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts in main clauses without inversion (SVO).

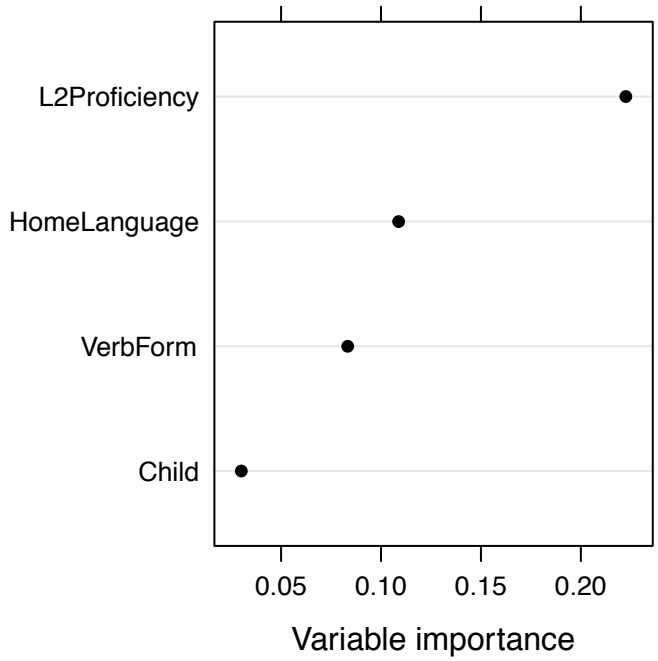
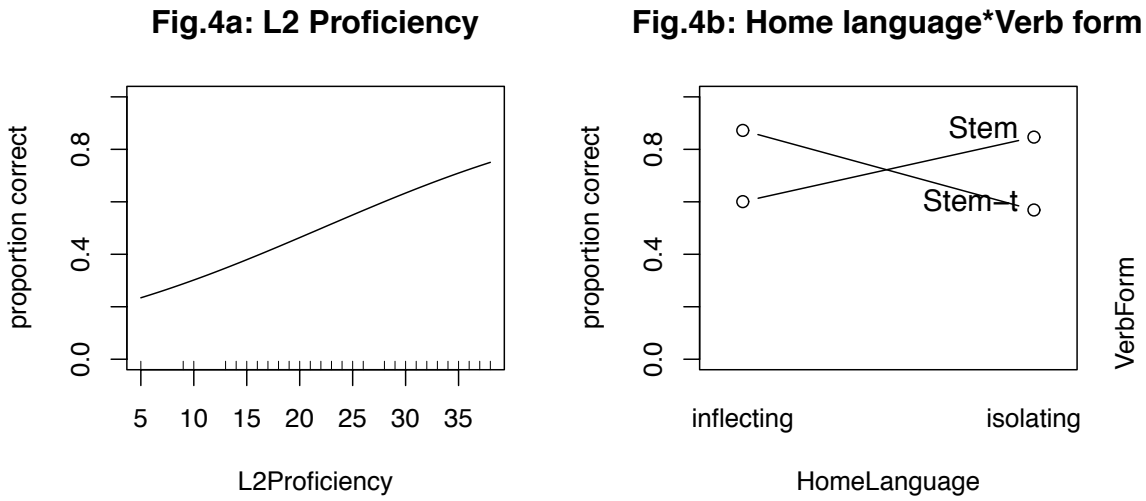


Figure 3: Variable importance plot of predictors (random forest); accuracy of subject-verb agreement in 1SG, 2SG, 3SG, 1PL and 3PL contexts in main clauses without inversion (SVO).



Figures 4a-b: Partial effects model 1 (mixed logistic regression); accuracy with subject-verb agreement in 2SG and 3SG contexts in main clauses with inversion (AdvVS) and without inversion (SVO).

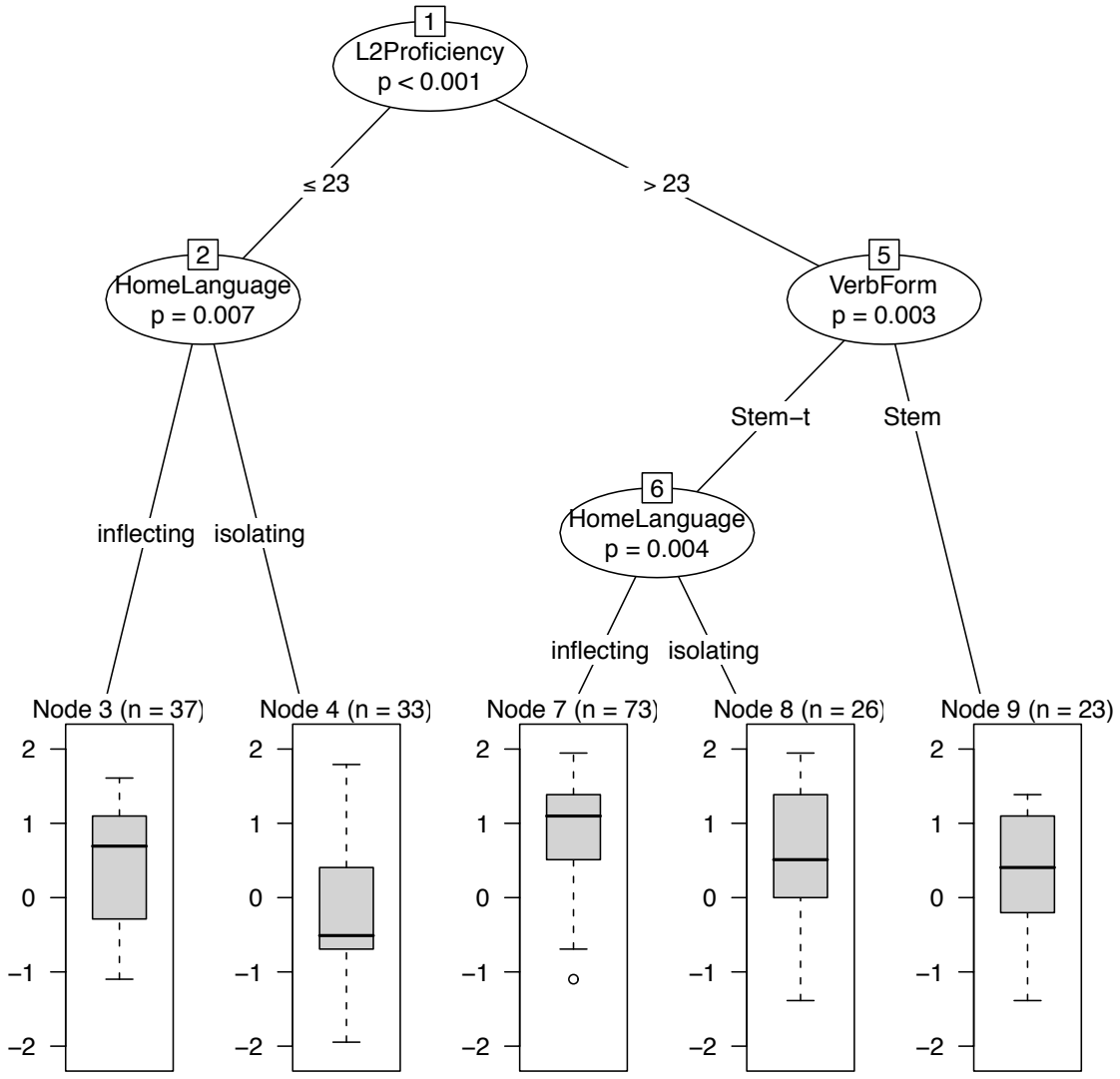


Figure 5: Conditional inference tree, predicting the empirical log-odds ratios; accuracy with subject-verb agreement in 2SG and 3SG contexts in main clauses with inversion (AdvVS) and without inversion (SVO).

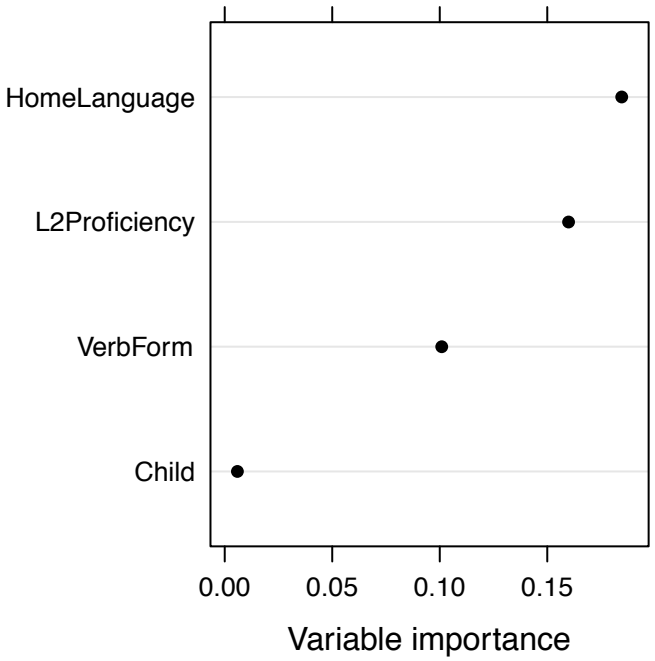
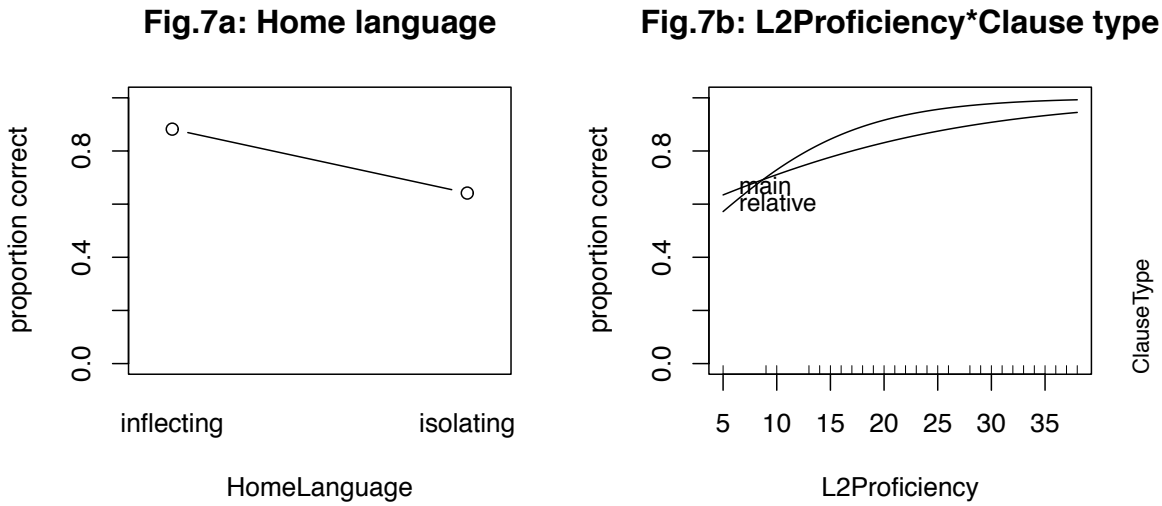


Figure 6: Variable importance plot of predictors (random forest); accuracy with subject-verb agreement in 2SG and 3SG contexts in main clauses with inversion (AdvVS) and without inversion (SVO).



Figures 7a-b: Partial effects model 1 (mixed logistic regression); accuracy with subject-verb agreement in 3SG contexts in main clauses without inversion (SVO) and relative clauses (SOV).

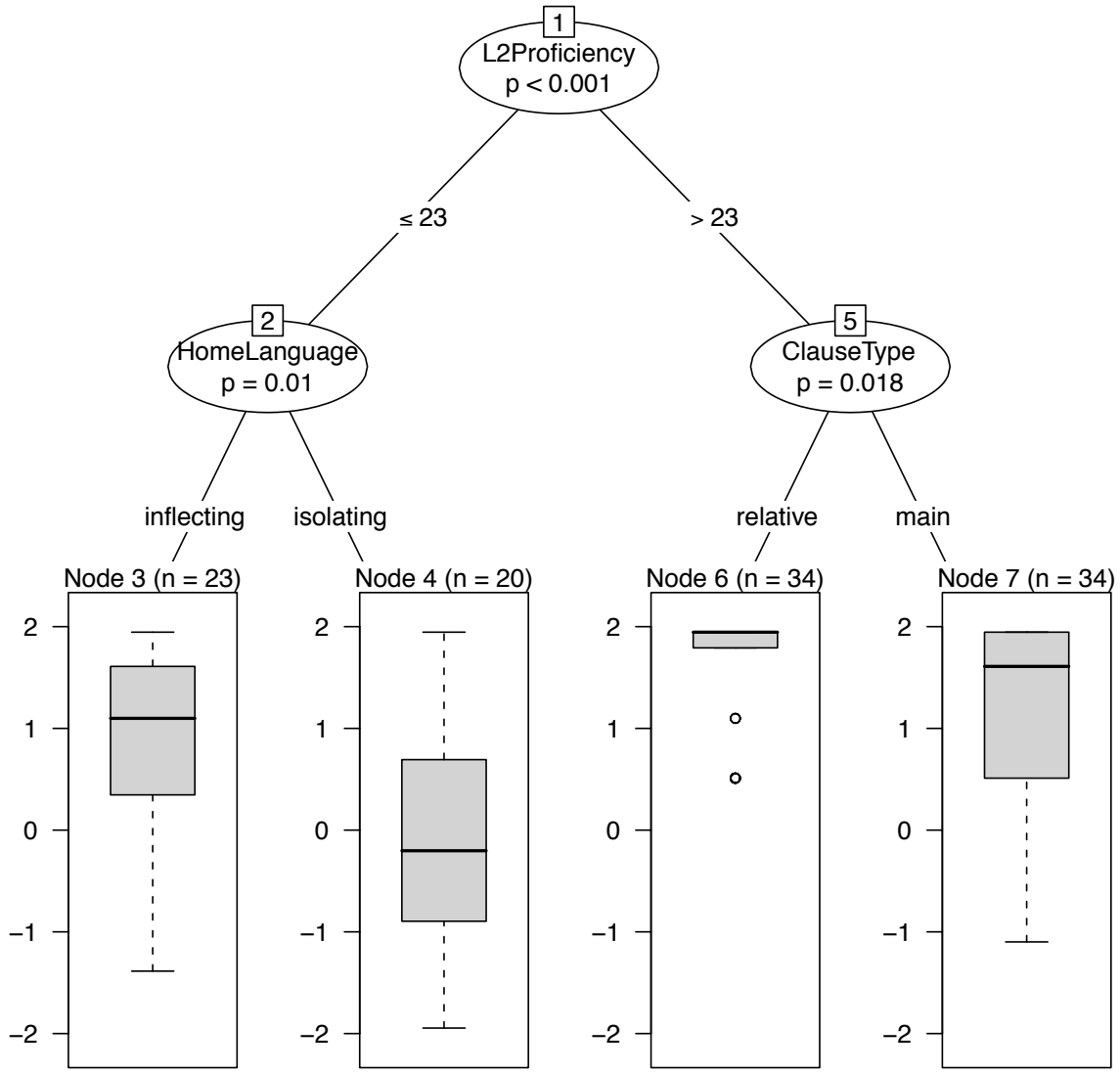


Figure 8: Conditional inference tree, predicting the empirical log-odds ratios; accuracy with subject-verb agreement in 3SG contexts in main clauses without inversion (SVO) and relative clauses (SOV).

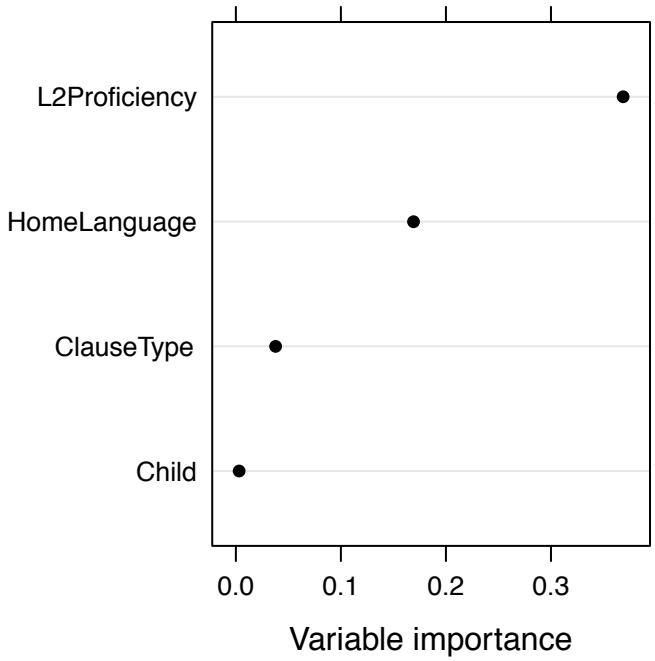


Figure 9: Variable importance plot of predictors (random forest); accuracy with subject-verb agreement in 3SG contexts in main clauses without inversion (SVO) and relative clauses (SOV).