LUND UNIVERSITY

The influence of lexical aspect and input frequency in the L2 French of adult beginners

Thomas, Anita

Link to publication

Total number of authors:
1

The influence of lexical aspect and input frequency in the L2 French of adult beginners

Anita Thomas

This article presents a study of the variation found in the oral production of verbs in L2 French by adult beginners with Swedish as L1. The study deals with the production of the two main forms of regular verbs in spoken French, a short form /parl/ and a long form /parle/, in present-tense and infinitive contexts. Learners at beginner stages use these two forms invariantly in both finite and non-finite contexts. The aim of this study is to investigate the choice of the invariant forms made by the learners. This is done by testing the influence of lexical aspect and input frequency. The investigation is made with concrete data from input sources. The results from both a free production task and an imitation test suggest that input frequency contributes more significantly than lexical aspect, although both factors overlap for most of the studied verbs.

Anita Thomas, Centre for Languages and Literature, Lund University, Box 201, SE-221 00 Lund, Sweden. anita.thomas@rom.lu.se

1. INTRODUCTION

The present study investigates the variation found in the oral production of verbs by adult Swedish learners of L2 French at beginner stages. More specifically, this study deals with the production of the two main forms of regular verbs in spoken French: a short form e.g. *parle* /parl/ 'speak' and a long form *parlE* /parle/ 'to speak' (among other meanings, see below for more details about this form), in present-tense and infinitive contexts. Table 1 illustrates the variation in the use of these two forms with four utterances produced by one of the learners in the study (age of onset of acquisition around 16 years) in a free production task. An asterisk indicates that the form is incorrect whereas parentheses indicate non-audible plural morphology.

The variation in the correct ((a) and (d) in Table 1) and incorrect ((b) and (c) in Table 1) production of verbs has been studied previously, within different theoretical frameworks. Most of these studies have focused on the interaction between the forms and the syntactic contexts into which they are inserted. Studies within the functional approach (Klein & Perdue 1997; Perdue, Benazzo & Guiliano 2002) have suggested that this variation is typical of learners at the stage they called 'the basic variety',

	Present-tense context	Infinitive context
Short form	a. Elle(s) mange(nt) en un restaurant. *they eat in a restaurant*	b. Et les hommes va *parle avec *and the men will speak with* les filles . . . *the girls*
Long form	c. Et il(s) *dansE ensemble. *and they dance together*	d. C'est deux personnes que va *it's two people who will* voyagE à en Italie. *travel to in Italy*

Table 1. Examples of variation in L2 French (free production task).

where 'there is no trace of FINITE verbs, in whatever function' (Klein & Perdue 1997:314; original emphasis). The forms produced by the learners at this stage correspond to base forms – infinitives or stems. Accordingly, temporal and other discourse relationships are expressed by other means (Klein & Perdue 1997:320). In their description of morpho-syntactic development in L2 French, Bartning & Schlyter (2004:285) came to a similar conclusion (based on two corpora of longitudinal data from Swedish-speaking learners), suggesting that the forms produced by learners at beginner stages are invariant. The use of short finite forms in finite contexts and long non-finite forms in non-finite contexts is seen as resulting from a gradual development towards target use of the language. On the other hand, studies within the generative framework, and especially within the Missing Surface Inflection Hypothesis (Prévost & White 2000), have suggested that syntax and morphology develop independently of each other in adult L2 learning, resulting in mismatches at surface level. These studies conclude that the use of non-finite forms in finite contexts 'are not in fact non-finite but are being used as default forms exhibiting properties of finite verbs' (Prévost & White 2000:109; see also Herschensohn 2001:297). The use of finite forms in non-finite contexts (as in (b) in Table 1) appears to be less frequent in L2 learners' utterances. Such forms have been examined by Prévost, who suggests that the occurrence of this error 'is due to L1 influence' (Prévost 2004:171). In sum, these generative accounts of the variation suggest that L2 learners of French at early stages of acquisition have some difficulties in providing the expected forms, but that the difficulties are not at the level of syntax (Prévost 2008:373).

From these previous studies I conclude that it may be inappropriate to reason in terms of finite and non-finite forms when referring to verb forms produced by L2 beginners. Indeed, these forms appear to behave much like default forms, i.e. as invariant forms used for different syntactic functions, in both finite and non-finite contexts, as was illustrated in Table 1. The question is whether there is any systematicity in L2 learners' selection of invariant forms, i.e. whether it is possible to predict the default form for a given verb. The aim of this study is to examine the possible influence of two factors, lexical aspect and frequency of forms in the input.

	Short form *parle* /parl/	Long form *parlE* /parle/
Stem	parl-	
Present tense	subject + parle/s/nt	vous parlez
Infinitive		parler
Past participle		parlé
Imperfect		subject + parlais/t/nt

Table 2. **The short and long form of the French regular verbs.**

These two factors have often been presented as being closely related to each other. At the same time only one of these factors has been said to influence beginners' production of verbs, viz. lexical aspect.

2. THEORETICAL BACKGROUND

2.1 Regular verbs

As mentioned above, the possible influence of lexical aspect and input frequency will be investigated concerning data from the class of regular verbs (*-er* verbs). This open class of verbs accounts for about 90% of French verbs, comprising approximately 4000 different verbs (Riegel, Pellat & Rioul 1994:263). The stem, e.g. *parl-* for the verb *parler* 'to speak', is homophonous with the forms of the present tense singular (*je parle* 'I speak', *tu parles* 'you speak', *il/elle parle* 'he/she speaks') and 3rd person plural (*ils/elles parlent* 'they speak'), all pronounced /parl/. A characteristic of French spoken language as compared to the other Romance languages is that the infinitive (*parler* 'to speak'), the past participle (*parlé* 'spoken'), the 2nd person plural present tense (*vous parlez* 'you.PL speak') as well as the imperfect singular and 3rd person plural (*je parlais*, *tu parlais* ... 'I spoke, you spoke ...') are all homophonous, i.e. pronounced /parle/ and usually orthographically transcribed *parlE*, where the *-E* corresponds to the sounds [e] and [ɛ]. The multi-functionality of the form *parlE* is one of the main characteristics of this verb class and constitutes a difficulty not only for L2 learners but also for native speakers when writing (e.g. Jaffré & Brissaud 2006). Following the terminology proposed by Bartning & Schlyter (2004), I will henceforth refer to these two forms as the short form (for *parle* /parl/) and the long form (for *parlE* /parle/) given that the long form also corresponds to finite functions. The functions of these two forms are presented in Table 2.

Taken together, these two forms cover the most frequent uses of this class of verbs in the spoken language, including the imperative and the subjunctive, neither of which will be further examined in this study.[1] Verb forms not covered by the long and the short forms are the conditional (*je parlerais* 'I would speak'), the morphological future (*je parlerai* 'I will speak') and the 1st person plural present tense (*nous parlons*

	States	Activities	Accomplishments	Achievements
Dynamicity	—	+	+	+
Durativity	+	+	+	—
Telicity	—	—	+	+
Examples	have, love	work, walk	write a book, grow	fall, find something

Table 3. Lexical aspect (classification based on Vendler 1957; Bardovi-Harlig 2000).

'we speak'). However, it should be noted that in spoken language, the 1st person plural is often replaced by the short form *on parle* 'one speaks' (3rd person singular).

As illustrated in Table 2, there is a close relationship between the short form and the present-tense form. It is thus plausible to suppose that L2 learners associate the present tense with the short form on the one hand, and, on the other hand, the long form with non-present-tense reference, past-tense (with or without an auxiliary), or infinitive in the periphrastic future (modal + long form).

2.2 Lexical aspect

A good starting point for tackling the question of systematicity in L2 learners' production is found in previous studies on the marking of temporality and past-tense contexts. The study of temporality within the European Science Foundation project, based on data from L2 learners in naturalistic contexts with different source and target languages (see e.g. Klein & Perdue 1997), suggested that beginners rely on lexical aspect (*Aktionsart*) to express temporal distinctions (Noyau et al. 1995; Perdue et al. 2002:860). The learners of French used mainly two basic forms, a V-Ø form and a V-[e] form (Noyau et al. 1995:167, her transcription), which correspond to what I refer to as short and long forms, respectively (see above). Their data showed the following distribution of these forms: the short form was used to express atelic situations and present-tense contexts and the long one for telic situations and past-tense contexts. It was only after some exposure to French that the forms alternated for the same verb (Noyau et al. 1995:160–175).

The influence of lexical aspect has largely been studied within the Aspect Hypothesis (Andersen 1991, 2002). This hypothesis is based on Vendler's (1957) semantic categories, which describe four situation types based on three main features: dynamicity, durativity and telicity (Vendler 1957; Bardovi-Harlig 2000:213f.). These categories are presented in Table 3.

The Aspect Hypothesis (Andersen 2002:79) makes the following predictions:

- past-tense marking will first appear on telic verbs and only later on atelic verbs;
- imperfective marking of past will appear later than perfective marking;
- imperfective marking will first appear with state and activity verbs and later with telic verbs.

According to Andersen, this development is due to cognitive principles that guide the learners' search for the appropriate forms. These principles can be summarised as follows (based on Andersen 2002:90–91):

- Learners tend to associate one form with a single meaning following a one-to-one principle of simplification of the target language.
- The prototypical meaning of a verb is more easily and naturally accessed than the less prototypical one.
- The first markers used by learners are those with the highest relevance to the meaning of the verb.
- From the markers available in the input, learners choose the marker most congruent with the meaning of the verb, for example, present tense and imperfect with state verbs.

Nonetheless, results from studies of L2 learners of French did not fully support the Aspect Hypothesis (see Thomas 2009 for an overview). Firstly, as predicted by the Aspect Hypothesis, learners produced the perfective, the French *passé composé* (*j'ai parlé* 'I have spoken', auxiliary + long form), earlier than the imperfective. However, this use was not restricted to telic verbs, but appeared with all dynamic verbs from the start. Learners may indeed see the *passé composé* as the past tense *par excellence*, as suggested by Kihlstedt (1998:108), who studied Swedish learners of French at advanced stages. Secondly, the use of the imperfective past, the French *imparfait* (*je parlais* 'I spoke'), was mostly restricted to state verbs without being extended to other verbs. In fact, even advanced learners preferred to use the present-tense form instead of the imperfect, especially with state verbs.

The results of these studies suggest that, when referring to PAST contexts, L2 learners of French tend to associate the present-tense form (short form) with state verbs and the *passé composé* (long form) with dynamic verbs. However, seeing that 'in the beginning stages of language acquisition ONLY INHERENT ASPECTUAL distinctions are encoded by verbal morphology, not tense or grammatical aspect' (Andersen 1991:307; original emphasis), the opposition between state and dynamic verbs might also be relevant for the study of the influence of lexical aspect in non-past-tense contexts, such as present tense and infinitive. Consequently, based on a hypothesis concerning the influence of lexical aspect, henceforth Hypothesis 1, learners at beginner stages are expected to use mainly the short form with state verbs and the long form with dynamic verbs, since the opposition between stative and dynamic matches the present-tense versus non-present-tense distribution of forms, as outlined in Section 2.1.

There is nonetheless an important question that has to be directed to the proponents of the influence of lexical aspect. This question concerns the relationship between the selection of the forms by the learners and the frequency of the forms in the input. Andersen's Distributional Bias Principle tries to capture this very close

relationship (cf. Andersen 2002:92f.), but in order to discuss the contribution of each of the factors, they need to be studied separately.

2.3 Frequency of forms in the input

The role of frequency in the input for L2 acquisition has mainly been considered within usage-based approaches, among others by Ellis (2002) and Bybee (2008). In these approaches 'grammar is viewed as the cognitive organization of one's experience with language' (Bybee 2008:216). The linguistic system is seen as dynamic, in the sense that it can be revised depending on experience and usage events, i.e. instances of production and comprehension of language (Kemmer & Barlow 2000:viii). This process is well summarised by Ellis & Cadierno (2009:112):

> Language structure emerges ontogenetically from usage in particular contexts. Development is slow and gradual, moving from an initial heavy reliance on concrete items to more abstract linguistic schema. This process is crucially dependent on the type and token frequencies with which particular constructions appear in the input.

Input can thus be defined as the language to which a learner is exposed (Kemmer & Barlow 2000:xv). In the case of an L2 learner, usages that are more or less target-like, including those expressed by other L2 learners and by the learner him/herself, might also be part of the input. Input and frequency of forms may thus vary depending on the context in which the learning takes place.

From the work of Bybee (1995:428f.; 2008:218f.) we know that type and token frequency plays a different role in establishing and maintaining memory representations. On the one hand, token frequency leads to higher autonomy and to weaker connections to other words. Moreover, it contributes to the word's 'lexical strength'; words with greater lexical strength will be more accessible. Type frequency, on the other hand, reinforces established connections resulting in patterns. These are more or less open and more or less productive depending on the level of type frequency, i.e. on the number of participants in the same pattern.

However, frequency alone is not sufficient for learners to notice and represent a form in long-term memory (cf. intake in Caroll 2001; VanPatten 2002). Learners obviously do not simply reproduce the language they are exposed to. For this reason input is also the language the learner actually processes. As pointed out by Schmidt (2001:7), one of the main problems of learners at beginner stages is that they are cognitively overloaded. Since working memory resources are limited, they have to be used in the best possible way. According to Schmidt, learners first direct their attentional resources 'at those elements that carry message meaning, primarily lexicon, and

only later, when the cost comes down, towards communicatively redundant formal features of language' (Schmidt 2001:13). This has also been underlined by VanPatten, who states that 'the intersection of high communicative value and frequency should have a favorable effect on acquisition' (VanPatten 2002:760).

In my understanding, the interaction between frequency and communicative value explains how the lexical verb can be salient in a construction with an auxiliary, whereas the auxiliary may go unnoticed.[2]

Based on this outline of the influence of input frequency on L2 acquisition, it seems reasonable to suggest that L2 beginners are sensitive mainly to token frequency. If these learners produce mainly one invariant form of a specific verb, we may expect it to be the most accessible form of each specific verb. Consequently, a second hypothesis, henceforth Hypothesis 2, can be made based on the interpretation of theoretical accounts of the influence of input frequency: learners are expected to produce the form that is the most frequent one in the input for each specific verb. That is, the learners are expected to produce the short form for the verbs with a high frequency of short forms in the input, and the long form for the verbs with a high frequency of long forms in the input. The specific predictions for the verbs will be based on a study of concrete input data (Section 4 below).

3. METHODOLOGY

3.1 Selection and classification of the verbs for this study

The present study is based on twelve verbs that L2 beginners are supposed to be familiar with. Some less frequent verbs have been added to meet the constraints of the tasks (number of state verbs, phonological factors).

In order to test Hypothesis 1 on lexical aspect, the verbs have to be allocated to the two categories discussed earlier, stative and dynamic, respectively. Such an allocation may be an impossible task, since many dynamic verbs can have a static reading (Vendler 1957:152). Based on the rationale behind the influence of lexical aspect, I have tried to find the PROTOTYPICAL lexical aspect of each verb by using a number of tests that have proven to work with French (see Table 4).

The final classification of the verbs is presented in (1) and (2) below. In order to test the two hypotheses of this study on a single verb, I added the verb *trouver* 'to find'. Just like its English equivalent, this verb has both a stative meaning as in 'to find the weather beautiful' and a dynamic meaning as in 'to find the keys'. From Hypothesis 1, the learners are expected to prefer the form that corresponds to the lexical aspect, i.e. the short form with the stative and the long form with the dynamic meaning. From Hypothesis 2, we expect a different pattern. The most frequent form of the verb in the input, this time taken to be a lexeme, should be the form that is

Test question	State verbs	Dynamic verbs
Is it possible to answer the question 'What happened?' by using the verb? (Vet 1980:55–56)	no	yes
Is the verb compatible with the question 'At what time?' (Vet 1980:55–56)	no	yes
Is the verb compatible with the adverb 'slowly'? (Platzack 1979:93–94)	no	yes
Is the verb compatible with the use of the French preposition 'de' in passive constructions? (Riegel et al. 1994:437)	yes	no

Table 4. **Tests for allocating verbs to categories (state versus dynamic).**

preferred by the learners. In other words, preference should be independent of lexical aspect.

(1) State verbs: *adorer* 'to adore', *aimer* 'to love', *détester* 'to detest', *habiter* 'to live (in a city, a flat)', *penser* 'to think', *préférer* 'to prefer', *trouver* 'to find something good/bad'.

(2) Dynamic verbs: *acheter* 'to buy', *manger* 'to eat', *parler* 'to speak', *regarder* 'to look at', *trouver* 'to find (after a search)', *visiter* 'to visit'.

As mentioned above, dynamic verbs like *parler* 'to speak' or *manger* 'to eat' have a static interpretation in some contexts, like in 'to speak a language' or 'to eat bread for breakfast'. Even if such uses may be very common in beginner classes, they do not correspond to the prototypical lexical aspect of the verbs.

3.2 Scores and statistics

The scores will be presented as tokens and proportions of the short form, because this form clearly corresponds to one single function. In order to distinguish the two hypotheses formulated for this study, the verbs will be investigated within the two categories of lexical aspect, states versus dynamic verbs for Hypothesis 1, and along a continuum of proportion of short forms for Hypothesis 2.

As the data is scarce, I will use non-parametric tools for statistical analysis. For Hypothesis 1, I will use the Mann-Whitney U test, which is similar to the t-test. Hypothesis 1 can be considered as verified if significantly more uses of the short form are observed with the stative verbs than with the dynamic ones. For Hypothesis 2, I will use the Spearman's rank-order correlation coefficient (Spearman's r) that describes the relationship between ranked scores. In this study I will compare ranked proportions of short forms. A correlation coefficient of 1 describes a perfectly linear relationship. Traditionally, a correlation of .30 is considered to describe a good correspondence between the scores.

Finally, the predictive contribution of the two factors will be tested with a logistic regression analysis.

3.3 Participants and tasks

The group of learners in this study consists of 33 secondary school students (average age 17.7 years) in Southern Sweden (Lund and Malmö). They have attended French classes at beginner stages for two to four terms. The learners' level of French will be established by analysing a free production task described below. The data were collected with one learner at a time and in the presence of the experimenter.

The control group consists of 10 native speakers of French coming from different parts of the world, six of whom were exchange students and four were living in Sweden.

The participants carried out three tasks: a free production task, a sentence imitation test (verbs in sentences) and a recall test (series of verbs). Due to space limitations, only the two first tasks will be presented in this paper.

4. STUDY OF INPUT DATA

The aim of this section is to examine input data in order to establish a continuum of frequencies for Hypothesis 2. I have tried to find input sources that are likely to reflect the main varieties of French to which L2 learners are exposed.

The first input source comes from two French-speaking interviewers in conversation with L2 learners at beginner stages. The data are taken from two corpora of free conversation: from my own corpus of classroom learners with 25 conversations each lasting 20–30 minutes and from four recordings of a classroom learner (Lisa) of around 60 minutes each in the Schlyter Lund Corpus. These data can be considered similar to conversations taking place in the classroom or with native speakers during stays in French-speaking countries. The learners exposed to this input source have been studied in detail in a pilot to this study (see Thomas 2009:105f.).

The second source of input comes from a corpus of classroom interactions, collected and analysed by Flyman Mattsson (2003). This corpus consists of 39 recordings of classroom teaching, with a total recording time of about 12 hours. In this source the verbs have been produced by secondary school teachers of French in classes at a somewhat higher proficiency level than the one of the learners in the present study. This input reflects the use of verbs by the teachers in activities such as grammar, translation, correction of exercises and general discussions.

The third source of input comes from the DVD C-ORAL-ROM (Cresti & Moneglia 2005), which comprises data of French as it is spoken in France. This is the kind of input the learners may find on holidays in France and also when teachers use authentic video and audio material in the classroom.

'to live (in)'	habite	habitais	habitez	habiter	habité
	108 (72%)	20 (13%)	15 (10%)	6 (4%)	1 (1%)
'to speak'	parle	parler	parlé	parlais	parlez
	220 (38%)	174 (30%)	92 (16%)	49 (8%)	46 (8%)
'to buy'	acheté	acheter	achète	achetez	achetais
	44 (40%)	38 (35%)	16 (15%)	10 (9%)	2 (2%)

Table 5. Frequency of all forms for three verbs (input).

In addition to these three input sources of spoken French, I have also looked at data from a written source, namely a Swedish textbook for beginners in French (*Escalade 1* by Tillman et al. 2001). The reason for incorporating these data was to have access to the verbs used in classroom activities at the very beginning of the learning process. The textbook *Escalade I* contains texts and exercises in a traditional way but also much material for communicative activities. Approximately fifty percent of the participants have used this book.

The ideal situation for studying the influence of input frequency would be to compare data from the teacher and the textbook that has been used with data from the learners in the same classroom, but this was unfortunately not possible this time. Nevertheless, the comparison of different input sources may also be seen as a strong point in this study as it might reveal similarities and differences in the use of language in the different sources.

To find out the most frequent form for each verb, all the forms presented in Table 2 above have been computed for all input sources. In a second step, the proportion of each form has been calculated from the total of forms found for each specific verb. These proportions have finally been ranked from the highest to the lowest, as illustrated in Table 5. The cells in grey indicate the short forms.

The results of this analysis showed that there are three groups of verbs with the same frequency pattern:

 (i) verbs that are very frequent in the short form, such as *habiter* 'to live (in)' in Table 5;
 (ii) verbs for which the short form is the most frequent in tokens but for which the long forms together are more frequent, such as *parler* 'to speak' in Table 5;
(iii) verbs where the long form is the most frequent and where the two most frequent forms are the two non-finite forms, such as *acheter* 'to buy' in Table 5.

The complete results for the twelve verbs and the four input sources are shown in Table 6. The table shows the proportion of short forms and, in parentheses, the sum of tokens (short + long forms). The verbs have been ranked from high to low, according to the proportion of short forms in the input as a whole ('sum input'). The dashed lines indicate the division into the three groups described above. It should be pointed out that for the verb *trouver* 'to find' the only interesting result here is the

	Interviewers	Teachers	C-ORAL	Textbook	Sum input
adorer 'to adore'	— (0)	— (0)	73% (11)	86% (21)	81% (32)
penser 'to think'	69% (29)	52% (31)	76% (307)	0% (1)	73% (368)
trouver (State)					
'to find something					
good/bad'	*57% (14)*	*40% (10)*	*75% (153)*	*100% (5)*	*72% (182)*
habiter 'to live (in)'	74% (34)	78% (18)	54% (54)	91% (44)	72% (150)
détester 'to detest'	— (0)	50% (2)	100% (4)	50% (4)	70% (10)
préférer 'to prefer'	0% (6)	63% (8)	63% (27)	100% (4)	58% (45)
aimer 'to love'	20% (104)	77% (26)	58% (142)	93% (69)	55% (341)
trouver (Total)					
'to find'	42% (19)	14% (51)	52% (286)	57% (37)	47% (393)
regarder 'to look at'	13% (39)	14% (37)	47% (128)	84% (19)	39% (223)
parler 'to speak'	29% (90)	38% (142)	36% (319)	80% (30)	38% (581)
manger 'to eat'	9% (11)	32% (75)	21% (62)	75% (12)	29% (160)
trouver (Dynamic)					
'to find (after a search)'	*0% (5)*	*7% (41)*	*26% (133)*	*50% (32)*	*26% (211)*
acheter 'to buy'	40% (5)	0% (19)	11% (72)	43% (14)	15% (110)
visiter 'to visit'	0% (8)	0% (2)	15% (13)	0% (2)	8% (25)

Table 6. Proportion of short forms (and sum tokens) in the different input sources.

total amount of tokens, i.e. *trouver* (Total). However, for the sake of completeness, the verb is also presented as *trouver* (State) and *trouver* (Dynamic).

The data presented in Table 6 show that there is some variation between the different input sources. Firstly, the data from the interviewers show a lower proportion of short forms than the other input sources. This result is due to a high proportion of 2nd person plural (e.g. *vous parlez* 'you speak') in the speech of one of the interviewers. This difference is above all visible for the verbs *préférer* and *aimer*, thus cutting down the proportion of short forms in the 'sum input'. Interestingly enough the pilot study showed that the learners of this interviewer never produced those verbs in any another form than the short form (see Thomas 2009:109). Secondly, the proportion of short forms in the textbook is high for all verbs except for the last two, *acheter* 'to buy' and *visiter* 'to visit'. This could mean that references to the here and now (the present) are the most frequent ones for beginners in a classroom situation. If this result reflects a general tendency in textbooks used today, the proportion of short forms may be high in the data from the participants in this study.

Despite the observed differences between the input sources, the correlation between them was over .35 (Spearman's r) except for the interviewers and the textbook (.21). However, the correlation between the different sources and the 'sum input' was

over .56, which means that the proportions presented under 'sum input' can be seen as representative for these input data.

We may now formulate Hypothesis 2, i.e. the one concerning the influence of the frequency of forms in the input:

(i) For the verbs that are very frequent in the short form in the input, we expect the learners to prefer the short form.

(ii) For the verbs where the short form is the most frequent in tokens but for which the long forms together are more frequent, we expect no real preference, i.e. a ratio closer to 50:50 than is the case for the other verbs.

(iii) For the verbs that are more frequent in the long form in the input, we expect the learners to prefer the long form.

Even if the proportion of short forms is significantly different across these three groups of verbs in the input (Mann-Whitney U, p < .01), recall that Hypothesis 2 will be tested by comparing the ranked proportions of short forms in the input and the learners' data. Nevertheless, such a grouping may facilitate the discussion of the learners' results.[3]

If we now compare the data in Table 6 with Hypothesis 1, which addresses the influence of lexical aspect, we can see that Hypothesis 1 and 2 overlap to a great extent. All the verbs that are frequent in the short form (from *adorer* to *aimer*) are indeed state verbs, and all the verbs that are more frequent in the long form (from *regarder* to *visiter*) are dynamic verbs. Consequently, there are only a few verbs that make it possible to separate the two hypotheses from each other, namely those dynamic verbs which are less clearly frequent in the long form in the input, *regarder*, *parler*, *manger* as well as *trouver* (Total) (the stative use of *trouver* is clearly frequent in the short form and the dynamic use is frequent in the long form).

5. RESULTS

5.1 Free production task

The 33 participants first completed an imitation test and then a free production task, but the data from the second task are presented first as they give some necessary indications of the learners' profiles.

The task was to tell a story from a set of pictures about two girls travelling to Italy (*Le voyage en Italie*, Ågren 2008). The instruction given was: 'You have 10 minutes to tell the story about what happens during the trip to Italy as if you were telling it to a person who can't see the pictures'. The learners were encouraged to tell the story in a descriptive way so as, in as far as this was possible, to avoid references to the past (indeed only one person made any references to the past). Interventions by the experimenter were minimal.

	Present-tense context	Infinitive context
Short form	157/256 (61%) correct	3/35 (9%) incorrect
Long form	99/256 (39%) incorrect	32/35 (91%) correct

Table 7. Correct and incorrect forms – free production task (33 learners).

With some exceptions, only the French words were transcribed (current words in other languages such as English *shopping*, *view* or Spanish *playa* 'beach' were transcribed, meta-comments in Swedish were not). The 33 learners produced 4790 words altogether (the ten native speakers produced 4049 words), but about a third of the learners produced less than 100 words each.

The linguistic proficiency level of the learners was established by using the L2 acquisitional stages for Swedish learners of French proposed by Bartning & Schlyter (2004). These stages are based on morpho-syntactic analyses of different linguistic phenomena such as finiteness, temporality, modality and aspect, negation, pronouns and subordination. Bartning & Schlyter propose six stages, from beginner (stage 1) to advanced (stages 4–6). The present study concerns learners at stages 1–3 (see Bartning & Schlyter 2004:294–295 for a detailed description). These stages correspond roughly to the 'pre-basic' and 'basic' varieties (see Klein & Perdue 1997).

The six stages have been translated into a computer program called *Direkt Profil* (Granfeldt et al. 2005), freely accessible on the Internet (http://profil.sol.lu.se:8080/profil/logon.jsp), which evaluates the proficiency level of a learner's text on the basis of 142 attributes, mainly measures of morpho-syntactic phenomena concerning the nominal and the verbal phrase, but also vocabulary measures. The result is given in a window showing the analysis for each attribute (this allows for a good opportunity to check how the transcription was analysed by the program) together with the level of proficiency in French. Even if this way of establishing the linguistic proficiency level of the learners has many weaknesses, it has the advantage of being less subjective than a more manual method.

The results obtained with *Direkt Profil* showed that none of the 33 learners had a higher level of proficiency in French than stage 3, which means that all the learners were at beginner stages (9 at stage 1, 14 at stage 2, and 10 at stage 3). The results of *Direkt Profil* were validated through the manual evaluation of six randomly chosen transcriptions rated by three expert judges.

All in all the learners produced 523 verb forms in the free production task. In order to gain an overview of the learners' linguistic profiles, the proportion of correct and incorrect forms in the context for the present tense and the infinitive (verbs in positions after a preposition or a modal verb) was computed for the 291 regular verb forms produced. The following forms did not enter into the calculation: 2nd person

plural (2 tokens), 1st person plural (3 tokens), forms ending in *-ante* and *-onte* (7 tokens, probably for 3rd person plural). Table 7 presents the tokens of forms in the same format as in Table 1 for all learners together since there is little difference between the learners grouped by stages for this task. The most important difference between the stages is that the context for the infinitive is not productive before stage 3 (these learners produced 10/14 infinitive contexts for regular verbs, 28/35 for the other lexical verbs). As expected for learners at beginner stages, there is a lot of variation in the data. For this reason the data are presented in raw scores.

The results of the correct and incorrect production of the verbs show:

- that the proportion of long forms in the context for the present tense is at least as high as expected for beginners according to earlier findings (Prévost & White 2000; Bartning & Schlyter 2004);
- that the proportion of errors in the context for the infinitive is lower than in the context for the present tense;
- that the context for the present tense is preferred by the learners of this study.

In sum, the data of these learners present nothing unexpected; they make errors, mostly in the context for the present tense but also in the context for the infinitive, as found in earlier studies (Bartning & Schlyter 2004; Prévost 2004).

We may now consider the production of the studied verbs in the free production task (*préférer* 'to prefer' and *détester* 'to detest' were not produced in this task). Table 8 presents the results for all the learners together. The verbs are ranked by the proportion of short forms in the learners' production (% short forms). The column 'Present-tense context/sum' shows the number of tokens produced in the context for the present tense and the sum of forms for the verb.

Table 8 shows that the five state verbs have all been produced in the short form at a rate of 80% or more, which was predicted by both hypotheses. The results for the dynamic verbs do not correspond to the prediction of Hypothesis 1, as only two verbs were clearly preferred in the long form, *acheter* and *visiter*. Instead, the results correspond to what is predicted by Hypothesis 2 concerning input frequency. The correlation between the ranked proportions of the short forms in learner and input data is .797 (Spearman's r), which describes a strong correlation. Such a correlation is an argument in favour of Hypothesis 2.

Furthermore, Table 8 shows again that the context for the present tense is the preferred context for these learners and for all the verbs. However, the results also show that the state verbs were never produced in a context other than the present tense. This raises the question as to how these verbs would be produced in the context for the infinitive. For this reason, an experimental imitation test was set up where the learners had to process all the verbs in both present-tense and infinitive constructions.

	Short forms %	Present-tense context/sum tokens	Lexical aspect	Input group
aimer' to love'	100%	20/20	S	1
habiter 'to live (in)'	100%	2/2	S	1
trouver (State) 'to find something good/bad'	*100%*	*2/2*	*S*	*—*
penser 'to think'	88%	17/17	S	1
adorer 'to adore'	80%	5/5	S	1
trouver (Total) 'to find'	67%	5/6	S+D	2
manger 'to eat'	63%	38/41	D	2
parler 'to speak'	61%	19/23	D	2
regarder 'to look at'	50%	15/16	D	2
trouver (Dynamic) 'to find (after a search)'	*50%*	*3/4*	*D*	*—*
acheter 'to buy'	27%	14/15	D	3
visiter 'to visit'	20%	4/5	D	3

S = stative verbs; D = dynamic verbs

Table 8. Proportion of short forms – free production task (33 learners).

5.2 Imitation test

5.2.1 Presentation

In an elicited imitation test, the learners hear a sentence which they have to repeat as exactly as possible. This test has been used for research on L1 and L2 language development (see Erlam 2006 for an overview). The principal advantage of this test is its proximity with the spoken production mode. It gives the possibility to force the encounter between the learner and structures s/he may try to avoid in free production (Naiman 1974:3). It involves the levels of comprehension, speech production, and memory capacity at the same time (Bley-Vroman & Chaudron 1994:247). When constructed in a meaningful way, it engages a RECONSTRUCTIVE process: 'In reconstruction, the subject hears the sentence and, to the extent that the sentence is comprehended, reconstructs the meaning using his or her own grammar' (Bley-Vroman & Chaudron 1994:246). For a sentence to be imitated correctly, it not only has to be understood, but the learners must also have enough grammatical knowledge and enough memory to re-encode the sentence heard. In other words, this test provides us with the possibility of comparing the grammar of the stimulus with the grammar of the learner.

In the construction of an elicited imitation test, three main factors have to be considered: the length and structure of the sentences, the grammatical level of the learner, and the memory capacity of the learner. To avoid rote imitation by the learner the sentence has to be long enough to exceed memory capacity but it must also be short enough to be processed by the learner.

5.2.2 Materials

In this study the material consisted of 27 sentences, two for each verb, one in the present tense, the short form, and another in the infinitive, the long form (see Thomas 2009:159–160). For the verb *trouver* 'to find' five sentences were created, two for each aspectual meaning and one sentence in the past-tense (*j'ai trouvé* 'I have found'). This sentence was introduced to test if this verb would be easier to remember in a sentence in the past-tense, and it was always presented as the very last test item to avoid all earlier reference to the past (the results of this sentence were disregarded in the end since past-tense is not studied here and since only a few learners could imitate it). Based on psycholinguistic findings for French (Segui & Ferrand 2002), the number of syllables was maintained as constant as possible, resulting in 19 sentences with 10 syllables, and 8 sentences with 11 syllables. The sentences were comprised of 5–10 words (average 6.85, s.d. 1.03) and the average length time of a sentence was 3.14 seconds (s.d. 0.31). As illustrated in examples (3)–(9), every sentence consisted of the following five parts:

 (i) a beginning with non-past reference, e.g. 'now', 'in Paris', 'at work';
(ii) a nominal subject, e.g. 'Marie';
(iii) a modal verb for the sentences containing an infinitive, *va* 'will' which has an interpretation of future time reference, and *veut* 'wants';
(iv) the test verb;
 (v) an end, e.g. 'the bus', 'a dress', 'with her boss'.

The test verb was placed in the middle of the sentence to avoid the serial effect of primacy and recency (Naiman 1974:25). All the sentences were grammatical and they were constructed so as to sound as natural as possible without turning into a typical fixed expression (chunk) that could have been learnt in the classroom. Four training sentences were created with irregular verbs, following the same rules.

The sentences were recorded by a native speaker of French in an echo-free room and in digital format. They were presented to the learners in four different lists; two mirror lists A and B and two lists with inverse order of the lists A and B, the 13 last sentences coming first. As mentioned above, the last sentence was always the same.

An articulatory suppression task was inserted between the point at which the learner had heard the sentence and the point at which s/he should recall it to disrupt the work of the phonological loop. According to Baddeley's memory model (see e.g. Baddeley et al. 1998), the phonological loop plays a central role in maintaining temporary items in memory. One of the functions of the phonological loop could be to help learn new words (Baddeley et al. 1998:158). Its capacity seems to be around two seconds (Speciale, Ellis & Bywater 2004:294). The articulatory suppression task used in this experiment is described in the procedure section below.

Stimulus	Output	Code
Short form	Short form	S>S
Short form	Long form	S>L
Short form	Other verb or nothing	S>øtr
Long form	Long form	L>L
Long form	Short form	L>S
Long form	Other verb or nothing	L>øtr

Table 9. Coding for the imitation test.

Finally, a cue picture was shown to the learners after the articulatory suppression task, as a help for recalling the content. The choice of the pictures was controlled for bias, i.e. the pictures should not overtly point to lexical aspect or grammatical structure.

5.2.3 Procedures and instructions

The test was presented to the learners in the form of slides on a computer. The procedure was as follows:

1. An exclamation mark appeared in the middle of the screen for two seconds followed by the sound of the recorded sentence.
2. After the sound the numbers 1–2–3–4–5–6–7–8–9 appeared on the screen for four seconds. At that moment the learner had to count once from 1 to 9 in French (articulatory suppression task).
3. When the numbers had disappeared, a cue picture appeared on the screen together with a dashed line; this was the moment where the learner had to recall the sentence or what s/he remembered.
4. When a sentence was finished, the learner had to press the return button to go on to the next sentence (self paced).

Instructions were given in Swedish by the experimenter. The learner was asked to repeat the sentence s/he had heard and it was emphasised that: 'if you can't remember everything, just say the words you remember even if it's only part of a sentence'.

5.2.4 Coding

The data were transcribed in a spreadsheet application following the structure of the stimulus. The verb forms (in grey in (3)) were coded by comparing the form in the stimulus and in the output. Recall that both hypotheses predict preferences for verb forms regardless of the syntactic context. Table 9 presents the coding of the outputs and example (3) presents the stimulus-sentence number 22 followed by the corresponding output of six learners (P25, etc.).

(3)	Stimulus	22	A Nice	Carl	va	trouver	que la mer est belle.	
			in Nice	*Carl*	*will*	*find*	*that the sea is beautiful*	
STAGE	LEARNER							CODE
1	P25	22	—	—	—	—	—	L>øtr
1	P33	22	En			—		L>øtr
3	P18	22		Carl		écrit 'writes'	Carl écrit la mer . . .	L>øtr
2	P21	22	A Nice	Carl		visitE '~visit'	la mer	L>øtr
3	P26	22		Carl		trouve	qui elle est belle	L>S
2	P19	22		Carl	va	trouver	la mer est belle	L>L

5.2.5 Results of the imitation test

All in all the learners processed 889 stimulus sentences (two subjects missed one sentence each). Three hundred and ten outputs had no verbs (see P25 and P33 in (3)), of which 52 had no words at all (P25 in (3)), and 143 outputs contained a verb other than the verb in the stimulus (P18 and P21 in (3)). The number of outputs without verbs decreased significantly from stage 1 to stage 3 (Mann-Whitney U, p < .05). This result is a way of validating the construction of the test, which is expected to be more difficult at lower L2 levels. In this study, the hypotheses will be discussed on the basis of the data from the 435 sentences (49%), i.e. only the outputs containing the same verb as in the stimulus (P26 and P19 in (3)), in accordance with the rationale behind the test.

For the verification of Hypothesis 1, Figure 1 shows the proportion of short forms for the stative and dynamic verbs for each of the three stages of the beginner-level proficiency in French, as well as for the control group of native speakers (see Table 10 below for the raw figures). Recall that all of the verbs were presented an equal number of times. The scores of the control group correspond exactly to the stimulus, i.e. they produced 50% short and 50% long forms for all the verbs. On the other hand, we can observe that the learners, unlike the native speakers, did not always repeat the verbs in the same form as was given in the stimuli. The state verbs were produced in the short form about 76% of the time, and even if there was some variation (see Table 10), this result corresponds to the prediction of Hypothesis 1 on lexical aspect. With respect to the production of the dynamic verbs, the proportion of short forms was higher than predicted, even though this proportion was lower than for the state verbs. The learners did not prefer the long form with the dynamic verbs contrary to what was expected from Hypothesis 1. In fact there were only two dynamic verbs that were preferred in the long form at all stages, namely the verbs *acheter* 'to buy' and *visiter* 'to visit', as in the free production task.[4]

These results suggest a limited influence of lexical aspect on the production of verbs by the learners. The result of the statistical analysis presents no significant difference for the proportion of short forms between the state and the dynamic verbs at stages 1 and 2 (p > .05). However, there is a significant difference at stage 3 (Mann-Whitney U, p = .038). This result is interesting, as it suggests that learners
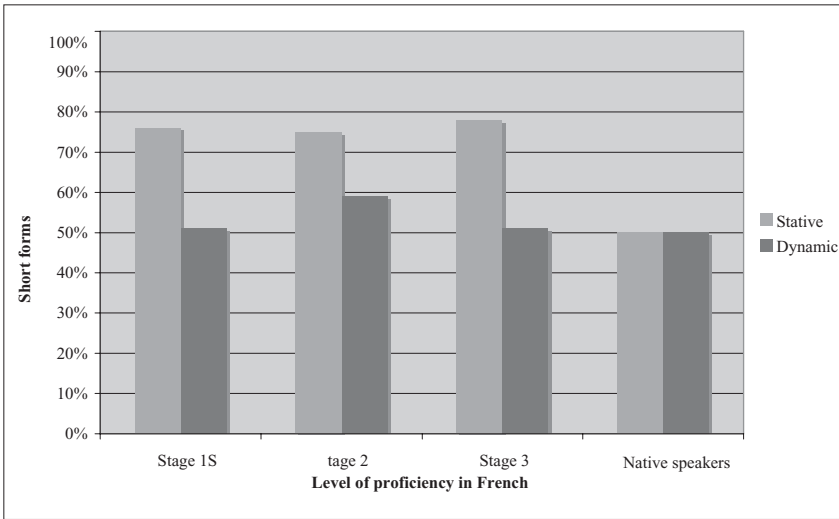
Figure 1. Proportion of short forms of the stative and dynamic verbs – imitation test.

may need a certain amount of experience of the target language before they produce the forms that correspond to the prototypical lexical aspect of the verbs. For the learners in this study, this seems to happen when they start producing contexts for the infinitive (see Section 5.1. above). At that time, the learners may have had enough experience to reproduce the typical patterns of association between form and lexical aspect. This could be explained as an effect of type frequency (Bybee 2008).

Turning now to Hypothesis 2, Table 10 shows the results for the twelve verbs for the learners at the three stages of the beginner-level proficiency in French. The scores indicate the sum of forms produced (Sum forms) and the proportion of short forms (short forms %). The last columns show the input frequencies from Table 6 as well as the groups of verbs in the input (group 1 overlaps with state verbs and groups 2 and 3 with dynamic verbs). The verbs are ranked from the highest to the lowest proportion of short forms for each stage of L2 French proficiency.

On the whole, the ranks of frequencies for the verbs at the three stages are very similar to the rank in the input. The Spearman's correlation coefficient between the input and stage 2 is .636 and between the input and stage 3 it is .684. Stages 2 and 3 show a very good correlation with the input. On the other hand, the correlation between the input and stage 1 is .354, which is weaker. This lower correlation is principally due to the results for the verbs *manger* 'to eat' and *penser* 'to think', which were produced contrary to what was predicted by input frequency. Generally, the number of verbs produced at stage 1 is lower than at the other stages. This could be due to the learners at this stage being yet unfamiliar with some of the verbs used in the test.

In order to discuss the data in a qualitative way, I will subsequently present a summary of outputs for three different verbs: *habiter* 'to live (in)', *acheter* 'to buy'

									Sum input (from Table 6)		
Stage 1	Sum forms	Short forms %	Stage 2	Sum forms	Short forms %	Stage 3	Sum forms	Short forms %	Input group		Short forms %
adorer	1	100%	penser	4	100%	aimer	10	100%	1	adorer	81%
aimer	4	100%	aimer	9	89%	adorer	6	83%	1	penser	73%
habiter	7	100%	habiter	18	83%	penser	11	82%	1	habiter	72%
manger	8	100%	parler	17	82%	habiter	14	79%	1	détester	70%
préférer	13	69%	regarder	22	77%	manger	14	79%	1	préférer	58%
détester	6	67%	adorer	4	75%	trouver	22	73%	1	aimer	55%
regarder	12	67%	détester	14	71%	détester	16	69%	2	trouver	47%
trouver	2	50%	préférer	20	70%	regarder	16	69%	2	regarder	39%
visiter	7	43%	trouver	16	69%	préférer	17	59%	2	parler	38%
parler	9	22%	visiter	22	45%	parler	16	56%	2	manger	29%
penser	1	0%	manger	15	40%	acheter	14	29%	3	acheter	15%
acheter	6	0%	acheter	15	27%	visiter	18	28%	3	visiter	8%

Table 10. Proportion of short forms at the three beginner-level stages of proficiency in L2 French – imitation test.

and *parler* 'to speak'. These verbs have been chosen because of their relatively high number of outputs.

For the verb *habiter* 'to live (in)', a preference for the short form is predicted by both hypotheses. Examples (4) and (5) show some representative outputs for the learners, one for each stage, the same learner for both outputs.

(4)		Stimulus	9	En France	Carl		habite	au bord de la mer.
				in France	*Carl*		*lives*	*at the sea side*
	STAGE	LEARNER						CODE
	1	P15	9		Carl		habite	près du mer	S>S
	2	P24	9	A la France	Carl		habite	près de la mer	S>S
	3	P06	9	En	Carl		habite	bord de la mer	S>S

(5)		Stimulus	10	En mars	Marie	veut	habiter	à Paris.
				in March	*Marie*	*will*	*live*	*in Paris*
	STAGE	LEARNER						CODE
	1	P15	10	En mars	någon		habite	de Paris	L>S
	2	P24	10	A mars	Carl		habite	à Paris	L>S
	3	P06	10	En mars	Carine	va	habite	à Paris	L>S

We can see that the short form was produced by all the learners, regardless of the stimulus. This result corresponds to the predictions of both hypotheses. The stimulus-sentence in the infinitive (5) was changed into a sentence in the present tense at stages 1 and 2. The learner at stage 3 was able to reproduce the syntactic structure of the stimulus, but she used the short form.

Examples (6) and (7) show some representative outputs for the verb *acheter* 'to buy', which is predicted to be produced in the long form by both hypotheses.

(6)		Stimulus	2	Aujourd'hui	Christine	veut	acheter	une robe.
				today	*Christine*	*wants*	*to buy*	*a dress*
	STAGE	LEARNER						CODE
	1	P10	2		Christi		achetE	un robe	L>L
	2	P16	2	Aujourd'hui	Marie		achetE	une robe	L>L
	3	P08	2		Catherine	va	acheter	une robe	L>L

(7)		Stimulus	1	Au supermarché	Carl		achète	un melon.
				in the	*Carl*		*buys*	*a melon*
				supermarket				
	STAGE	LEARNER						CODE
	1	P10	1	En supermarché			achetE	un	S>L
	2	P16	1	A supermarché	Carl		achetE	un melon	S>L
	3	P08	1	A supermarché	Carl		achète	un melon	S>S

For this verb, we can see that the learners at the first two stages systematically produced the long form but without marking the syntactic difference. In other words, they produced the construction of the present tense, subject + verb, a construction

they are very familiar with (cf. free production data in Section 5.1 and Table 7 above). The learner at stage 3, on the other hand, was able to correctly imitate the structure of the infinitive with the predicted form, and she also correctly imitated the stimulus in the present tense. The results for this verb also lend support to both hypotheses.

Finally, examples (8) and (9) show some representative outputs for the verb *parler* 'to speak', which is predicted to be produced in the long form by Hypothesis 1 as it is a dynamic verb, and in both forms by Hypothesis 2.

(8)   Stimulus 14 Maintenant Marie veut parler avec son chef.
                             now   *Marie*   *wants*   *to.speak*   *with her boss*

STAGE	LEARNER						CODE
1	P31	14				parlE	en su chef	L>L
2	P19	14 Maintenant Marie			parle	avec son chef	L>S
3	P26	14 Maintenant Marie	va	parle	avec son chef	L>S
3	P08	14 Maintenant Carine	va	parler	avec son chef	L>L

(9)   Stimulus 13 Au travail Carl parle avec une collègue
                             at work   *Carl*   *speaks*   *with a*   *colleague*

STAGE	LEARNER				CODE
1	P31	13		parlE	dans une copain	S>L
2	P19	13 Au travail Carl	parle	avec une collègue	S>S
3	P26	13 Au travail Carl	parle	au une collègue	S>S
3	P08	13 Au travail Carl	parle	avec une collègue	S>S

In general, we can see that there is more variation in the production of this verb than for the other two, *habiter* 'to live (in)' and *acheter* 'to buy' presented above. We can see that the learner (P31) at stage 1 produced the verb in the long form, which could be an argument in favour of the influence of lexical aspect. The fact that the verb is not preceded by a subject gives the impression of a lexical chunk being used, 'speak to a colleague/speak to the boss'. The learner (P19) at stage 2 only produced the verb in the short form and so did most of the learners at this stage. Finally, at stage 3, we find two types of learners: those like P26, producing this verb in the short form even within the infinitive construction and those like P08, imitating the form and construction of the stimulus. The fact that we find some more variation in the production of this verb can be seen as an argument for the influence of the forms in the input, despite the fact that the learners at stage 1 preferred the long form for this verb.

A logistic regression analysis was conducted, including all predictors, i.e. the lexical aspect (stative or dynamic), the relative frequency in the input (in percentage), and the form presented in the stimulus (present tense or infinitive). In a model with all three predictors together, the predictor 'lexical aspect' was not significant, but 'input frequency' and 'stimulus type' were. 'Input frequency' was still significant ($p > .001$) when it was introduced after (not simultaneously with) 'stimulus type' (this predictor appeared to be quite strong) (for details, see Thomas & van de Weijer 2010). This

means that input frequency but not lexical aspect contributed significantly to the results of the L2 learners in the imitation test. These results confirm the tendencies found in the qualitative analysis.

6. DISCUSSION

The aim of this study was to investigate the variation of verbs in oral production by adult Swedish learners of L2 French at beginner stages. Earlier studies (Klein & Perdue 1997; Bartning & Schlyter 2004; Prévost 2004) have suggested that the verb forms produced by beginners resemble mainly invariant forms, used both in finite and non-finite contexts (see Table 1). This raised the question whether we can observe some systematicity in the learners' production of forms. For this purpose I looked at two factors, lexical aspect and input frequency, in the production of twelve regular *-er* verbs in French. From the presentation of these verbs in Section 2.1 it appeared that the two main forms in spoken French, the short form e.g. /parl/ 'speak' and the long form e.g. /parle/ mainly correspond to a distribution of present-tense versus non-present-tense.

Based on earlier studies on the influence of lexical aspect in L2 French (Noyau et al. 1995; Andersen 2002), a first hypothesis, Hypothesis 1, was made. According to this hypothesis, learners at beginner stages would mainly use the short form with state verbs and the long form with dynamic verbs, regardless of the context, present tense or infinitive.

A recurring discussion in studies on the influence of lexical aspect has focused on the role of input frequency in the learners' selection of forms. In the present study, I therefore tried to investigate each factor separately. The second hypothesis, Hypothesis 2, was thus based on earlier accounts of the influence of frequency of forms in second language learning (Ellis 2002; Bybee 2008). According to this hypothesis, the learners were expected to produce the most accessible form for a specific verb, namely the most frequent form of this specific verb in the input.

A first empirical step involved looking at different input sources in order to establish the most frequent form for all the verbs included in the study. Based on the proportion of short forms, the following three groups of verbs emerged:

 (i)   verbs frequent in the short form (i.e. the present tense);
 (ii) verbs frequent in both short and long forms at a level of about 50% for each form;
(iii) verbs frequent in the long form.

It turned out that the predictions of the lexical aspect hypothesis (Hypothesis 1) and input frequency hypothesis (Hypothesis 2) overlapped for the verbs in the first group (state verbs) and the third group (dynamic verbs; see Table 6 and Section 4 above). The only way to have the two hypotheses make clearly distinctive predictions was to look at a group of dynamic verbs with the same input frequency with regard

to the two forms (second group). Thus, the results of the input study confirmed the close relationship between lexical aspect and input frequency, as well as the necessity to investigate both factors independently.

The two predictions were then examined in data from 33 learners of French at beginner stages. This was done in a free production task and in an imitation test. A brief analysis of the learners' correct and incorrect production of the two studied forms in present-tense and infinitive contexts showed the usual proportion of errors for learners at these stages (Bartning & Schlyter 2004; Prévost 2004).

The first examination of the twelve verbs studied in the free production task showed a strong correlation (r = .797) between the ranked proportions of short forms in the input and in the data from the learners. However, the learners never produced state verbs in any context other than the one of the present tense. This made it difficult to observe the use of an invariant form, as the learners did not create contexts for the use of a form other than the short form.

In the imitation test, all verbs had to be processed by the participants, once in the short form in a construction of the present tense and once in the long form in a construction for the infinitive (modal + long form). The results can be summarised as follows:

- In general, there was a preference for the use of the short form for all verbs except the two verbs that were most frequent in the long form in the input (*acheter* 'to buy' and *visiter* 'to visit'). This preference for the short form can be seen as reflecting the frequency of forms in the input as represented by the textbook (see Table 6). This preference also suggests that learners' production is influenced by specific classroom usage of forms (Ellis 2008).

- An influence of lexical aspect could only be observed for the slightly more advanced learners (stage 3 of Bartning & Schlyter 2004 classification of L2 proficiency levels) but not for those who were at the earlier stages of learning. This is contrary to what has been claimed by previous studies (Andersen 1991; Perdue et al. 2002) and suggests that the learners might need some experience of the target language before they can start producing the forms that correspond to the prototypical distribution found in the input.

- The correlation between the ranking of proportions of short forms in the input on the one hand and in the learners' data on the other turned out to be a good correlation. However, the correlation between the input and the production of the beginner learners at a very early stage (stage 1) was weaker. This suggests that the learners might need some experience before one form becomes more accessible than another.

- A qualitative discussion of some representative outputs showed differences in processing and producing the infinitive, dependent on the verb involved. Indeed, the infinitive construction was hardly used at all with verbs that were frequent in the short form in the input. This may explain why the incorrect use

of the short form is less frequent than the incorrect use of the long form in L2 learners French (see Section 5.1 and Table 7). Nevertheless, in the few cases where the infinitive construction was used, it was used with a short form, which is grammatically incorrect but expected from the influence of token frequency. On the other hand, the same construction turned out to be easier to imitate for verbs that were frequent in the long form in the input.

- Target-like production of the infinitive construction (modal + long form) appeared mainly at stage 3. This suggests that this might be a stage where some effects of type frequency can be observed. That would confirm Bybee's suggestion that 'a certain degree of type frequency is needed to uncover the structure of words and phrases' (Bybee 2008:221).
- Finally, a logistic regression analysis confirmed the predictive contribution of the frequency of the input studied here.

The results of the imitation test thus suggest an influence of input frequency on the beginners' production of invariant forms in present-tense and infinitive contexts. It would be interesting to compare these findings with similar studies of other linguistic phenomena. In addition, a study of the phonological characteristics of the forms would be an important follow-up to the present study.

ACKNOWLEDGEMENTS

NOTES

1. The forms in Table 2 cover about 90% of the forms of the verbs examined in this study, based on the tokens found on the DVD of French spoken language C-ORAL-ROM (Cresti & Moneglia 2005).
2. Indeed, very frequent items, such as grammatical markers, are often monosyllabic and unstressed (Labelle 2005:445), which makes them difficult to notice and, consequently, to acquire (Ellis 2008:381f.).
3. The fact that some verbs are more frequently used in the input than others, as illustrated in Table 6, can give us an indication of how familiar the learners might be with the different verbs. However, this factor will not be studied separately in this paper.
4. The data for the two meanings of the verb *trouver* 'to find' do not appear in Table 10. The stative meaning has been produced by the learners with 0% short forms at stage 1, 43% at stage 3 and 92% at stage 3, and the dynamic meaning with 100%, 89% and 50% short forms, respectively.

REFERENCES

Ågren, Malin. 2008. *À la recherche de la morphologie silencieuse. Sur le développement du pluriel en français L2 écrit*. Ph.D. dissertation, Lund University.

Andersen, Roger W. 1991. Developmental sequences: The emergence of aspect marking in second language acquisition. In Thom Huebner & Charles A. Ferguson (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*, 305–324. Amsterdam & Philadelphia: John Benjamins.

Andersen, Roger. 2002. The dimensions of 'pastness'. In Rafael Salaberry & Yasuhiro Shirai (eds.), *The L2 Acquisition of Tense–aspect Morphology*, 79–105. Amsterdam & Philadelphia: John Benjamins.

Baddeley, Alan D., Susan Gathercole & Costanza Papagno. 1998. The phonological loop as a language learning device. *Psychological Review* 105(1) 158–173.

Bardovi-Harlig, Kathleen. 2000. *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use*, special issue of *Language Learning* 50(s1).

Bartning, Inge & Suzanne Schlyter. 2004. Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies* 14 281–299.

Bley-Vroman, Robert & Craig Chaudron. 1994. Elicited imitation as a measure of second-language competence. In Elaine E. Tarone, Susan M. Gass & Andrew D. Cohen (eds.), *Research Methodology in Second-language Acquisition*, 245–261. Hillsdale, NJ: Lawrence Erlbaum.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5) 425–455.

Bybee, Joan. 2008. Usage-based grammar and second language acquisition. In Robinson & Ellis (eds.), 216–236.

Carroll, Susanne E. 2001. *Input and Evidence: The Raw Material of Second Language Acquisition*. Amsterdam & Philadelphia: John Benjamins.

Cresti, Emanuela & Massimo Moneglia (eds.). 2005. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam & Philadelphia: John Benjamins.

Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24 143–188.

Ellis, Nick C. 2008. Usage-based and form-focused language acquisition. In Robinson & Ellis (eds.), 372–405.

Ellis, Nick C. & Teresa Cadierno. 2009. Constructing a second language. *Annual Review of Cognitive Linguistics* 7 111–139.

Erlam, Rosemary. 2006. Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics* 27(3) 464–491.

Flyman Mattsson, Anna. 2003. *Teaching, Learning, and Student Ouput: A Study of French in the Classroom.* Ph.D. dissertation, Lund University.

Granfeldt, Jonas, Pierre Nugues, Emil Persson, Lisa Persson, Fabian Kostadinov, Malin Ågren & Suzanne Schlyter. 2005. Direkt profil: A system for evaluating texts of second language learners of French based on developmental sequences. *The Second Workshop on Building Educational Applications Using Natural Language Processing, 43rd Annual Meeting of the Association of Computational Linguistics*, 53–60. Ann Arbor, MI.

Herschensohn, Julia. 2001. Missing inflection in second language French: Accidental infinitives and other verbal deficits. *Second Language Research* 17(3) 273–305.

Jaffré, Jean-Pierre & Catherine Brissaud. (eds.). 2006. *Morphographie et homophones verbaux* (Langue Française 151). Paris: Larousse/Armand Colin.

Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based Models of Language*, vii–xxviii. Stanford, CA: CSLI Publications.

Kihlstedt, Maria. 1998. *La référence au passé dans le dialogue. Étude de l'acquisition de la temporalité chez des apprenants dits avancés de français.* Ph.D. dissertation, Stockholm University.

Klein, Wolfgang & Clive Perdue. 1997. The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13(4) 301–347.

Labelle, Marie. 2005. The acquisition of grammatical categories: The state of the art. In Henry Cohen & Claire Lefebvre (eds.), *Handbook of Categorization in Cognitive Science*, 433–457. Amsterdam: Elsevier.

Naiman, Neil. 1974. The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism* 2 1–37.

Noyau, Colette, Et-Tayeb Houdaïfa, Marie-Thérèse Vasseur & Daniel Véronique. 1995. The acquisition of French. In Rainer Dietrich, Wolfgang Klein & Colette Noyau (eds.), *The Acquisition of Temporality in a Second Language*, 145–209. Amsterdam & Philadelphia: John Benjamins.

Perdue, Clive, Sandra Benazzo & Patrizia Guiliano. 2002. When finiteness gets marked: The relation between morphosyntactic development and use of scopal items in adult language acquisition. *Linguistics* 40(4) 849–890.

Platzack, Christer. 1979. *The Semantic Interpretation of Aspect and Aktionsarten: A Study of Internal Time Reference in Swedish*. Dordrecht: Foris.

Prévost, Philippe. 2004. Morphological variation in early adult second language French: A cross-sectional study. In Susan H. Foster-Cohen, Michael Sharwood Smith, Antonella Sorace & Mitsuhiko Ota (eds.), *EUROSLA Yearbook 4*, 147–175. Amsterdam & Philadelphia: John Benjamins.

Prévost, Philippe. 2008. Knowledge of morphology and syntax in early adult L2 French: Evidence for the Missing Surface Inflection Hypothesis. In Juana Liceras, Helmut Zobl & Helen Goodluck (eds.), *The Role of Formal Features in Second Language Acquisition*, 352–377. New York: Lawrence Erlbaum.

Prévost, Philippe & Lydia White. 2000. Missing surface inflection or impairement in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16(2) 103–133.

Riegel, Martin, Jean-Christophe Pellat & René Rioul. 1994. *Grammaire méthodique du français*. Paris: PUF.

Robinson, Peter & Nick C. Ellis (eds.). 2008. *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge.

Schmidt, Richard. 2001. Attention. In Peter Robinson (ed.), *Cognition and Second Language Instruction*, 3–32. Cambridge: Cambridge University Press.

Segui, Juan & Ludovic Ferrand. 2002. The role of the syllable in speech perception and production. In Jacques Durand & Bernard Laks (eds.), *Phonetics, Phonology and Cognition*, 151–167. Oxford: Oxford University Press.

Speciale, Giovanna, Nick C. Ellis & Tracey Bywater. 2004. Phonological sequence learning and short-term memory store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics* 25 293–321.

Thomas, Anita. 2009. *Les apprenants parlent-ils à l'infinitif? Influence de l'input sur la production des verbes par des apprenants adultes du français* (Études romanes de Lund 87). Ph.D. dissertation, Lund University. http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1473797&fileOId=1473802 (accessed 16 August 2010).

Thomas, Anita & Joost van de Weijer. 2010. Regressionsanalys av L2 produktionsdata. Presented at Seminar on Language Learning, 27 April 2010, Lund University.

Tillman, Birgitta, Matts Winblad, Sylvia Martin, Véronique Lönnerblad, Marie Pettersson & Ewa Sandberg. 2001. *Escalade 1*. Stockholm: Liber.

VanPatten, Bill. 2002. Processing instruction: An update. *Language Learning* 52(4) 755–803.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2) 143–160.

Vet, Co. 1980. *Temps, aspects et adverbes de temps en français contemporain. Essai de sémantique formelle.* Genève: Droz.