

University of Groningen

Dynamic Usage-Based Principles in the Development of L2 Finnish Evaluative Constructions

Lesonen, Sirkku; Steinkrauss, Rasmus; Suni, Minna; Verspoor, Marjolijn

Published in:
Applied Linguistics

DOI:
[10.1093/applin/amaa030](https://doi.org/10.1093/applin/amaa030)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Lesonen, S., Steinkrauss, R., Suni, M., & Verspoor, M. (2021). Dynamic Usage-Based Principles in the Development of L2 Finnish Evaluative Constructions. *Applied Linguistics*, 42(3), 442-472.
<https://doi.org/10.1093/applin/amaa030>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Dynamic Usage-based Principles in the Development of L2 Finnish Evaluative Constructions

^{1,2,*}SIRKKU LESONEN, ³RASMUS STEINKRAUSS, ⁴MINNA SUNI, and ^{5,6}MARJOLIJN VERSPOOR

¹Department of Language and Communication Studies, University of Jyväskylä, Atheneum, Building A, Seminaarinkatu 15. PO Box 35 FI-40014, ²Faculty of Arts, The Center for Language and Cognition Groningen, University of Groningen, ³Faculty of Arts, Department of Applied Linguistics, University of Groningen, PO Box 716, 9700, AS Groningen, ⁴Department of Language and Communication Studies, University of Jyväskylä, Atheneum, Building A Seminaarinkatu 15, PO Box 35 FI-40014, ⁵Faculty of Arts, Department of Applied Linguistics, University of Groningen, PO Box 716 9700, AS Groningen and ⁶Department of Applied Linguistics Veszprem, University of Pannonia

*Corresponding author: Department of Language and Communication Studies, University of Jyväskylä, Atheneum, Building A, Seminaarinkatu 15 PO Box 35 FI-40014. E-mail: sirkku.lesonen@jyu.fi

This study investigates the formal verbalizations of evaluation used by four beginning L2 learners of Finnish from a dynamic usage-based perspective. Longitudinal data collected weekly were used to investigate what kind of constructions learners use to express evaluation and how these interact and develop over time. The results show that when a new construction is acquired in the L2, another related construction might regress. The results also point to increased variability in the construction during a phase of rapid development and reduced variability in the phases of regression or slower progress. These findings add to our understanding of a developing L2 as a system in which changes in one aspect have the potential to bring about changes in interconnected aspects. The variability patterns found in the learners' developmental trajectories add to the growing body of research that proposes variability as meaningful in the learning process.

INTRODUCTION

In recent years, researchers have increasingly come to view language development as a complex and dynamic process (de Bot *et al.* 2005; Larsen-Freeman and Cameron 2008; van Geert 2008; van Geert and Verspoor 2015). To get more insight into the actual developmental process, several longitudinal case studies have focused on non-linear L2 developmental patterns: variation among learners, the variability in developmental measures, and changing

interaction between subsystems in the learner's language. For example, it has been shown that developing complexity is a necessary precursor for accuracy in L2 English (Caspi 2010). Moreover, a competitive relationship between developing vocabulary and syntax has been found in L2 English learning (Verspoor *et al.* 2008). So far, studies within this framework have focused exclusively on the linguistic forms that were produced by tracing the development of the common complexity, accuracy, and fluency (CAF) measures. No attention has been paid to the other side of the coin: how do language learners, with their limited linguistic repertoire, manage to convey a certain meaning and how do the various linguistic means interact over time? The current study aims to fill this gap by taking meaning as the starting point.

This study will focus on how four beginning learners of L2 Finnish express evaluation, that is, that something is good or bad, desirable or non-desirable in their opinion, and investigate how different types of linguistic constructions expressing this notion develop and interact with each other over time. In an earlier single case study (Lesonen *et al.* 2017), we found that an L2 learner used two main types of constructions to express evaluation, namely verbal (e.g. *Minä tykkäsin *kaikki ruoasta* 'I liked all the food') and adjectival (e.g. **kasvustudeiden on tosi hyvä* 'the education (science) is very good') constructions. As in the previous Complex Dynamic Systems Theory (CDST) studies focusing on linguistic forms, we found a clear non-linear development with two constructions competing in the early stages: initially, the learner used verbal constructions almost exclusively, which developed at the expense of adjectival constructions. Later, when adjectival constructions were used more frequently, the frequency of verbal constructions leveled off.

The aim of this article is to examine to what extent such dynamic patterns can be found in other learners with different L1s in the same learning context. The article is organized as follows. After presenting a theoretical framework of dynamic usage-based (DUB) linguistics, we will review previous work on the interaction of different aspects in the developing L2 and on intra-individual variability in L2 development. The empirical study itself looks at the quantitative development of evaluative constructions in four beginner learners of Finnish.

L2 DEVELOPMENT FROM A DUB PERSPECTIVE

In the context of CDST and usage-based linguistics—a dynamic usage-based (DUB) perspective for short—it is assumed that usage events, that is, the use of language in particular contexts, drive L2 development and that the emergent grammar of the L2 evolves in discourse (Eskildsen 2009; Langacker 2009). A learner's language, therefore, changes every time the learner uses the language in social interaction. During the changes, the learner's language undergoes over time, the learner might move toward more effective ways of using the L2: the learner is progressing. Sometimes the changes might temporarily lead to less effective communication strategies and a regression. Both progress and regress result in variability and both can be seen as development (Larsen-Freeman 2013).

Several studies have shown that developmental trajectories in L2 learning can be quite different from one individual to another (Lowie and Verspoor 2015). Even identical twins, whose exposure to the target language is very similar, may show clear differences in their L2 learning trajectories (Chan *et al.* 2015). Also, the 22 L2 English learners in the same learning context traced over one academic year by Lowie and Verspoor (2019) had vastly different developmental paths in terms of lexical and syntactic complexity. As the authors point out, ‘no two individuals will develop in exactly the same manner as development takes place in a non-linear fashion, with phases of high degrees of variability accompanying rapid development’ (Lowie and Verspoor 2019: 185). Regardless of the highly individual learning trajectories, these two developmental principles mentioned by Lowie and Verspoor (2019), namely the non-linearity of development (the alternation of progress and regress) and the increased variability when progress is taking place, could possibly be generalized to a bigger learner group.

In a DUB approach, the non-linear development in L2 learning is assumed to be based on the self-organization of the different subsystems of the L2. Different subsystems of the L2, such as lexicon, phonology, and syntax (de Bot and Larsen-Freeman 2011) or CAF (Housen *et al.* 2012), interact with each other continuously. These subsystems also interact with the learner’s internal resources (e.g. motivation, aptitude, and attention) and external resources (the target language environment) (de Bot and Larsen-Freeman 2011). A change in one subsystem has the potential to affect other aspects and hence the whole system, and non-linear changes emerge from these interactions. The various subsystems may develop rather independently before they self-organize and become coordinated. The different subsystems of the L2 may therefore exhibit various kinds of interactions, and these may change over time as development takes place (van Geert 2008). The first type of interaction is a conditional relationship in which one subsystem has to reach a certain minimal level before another subsystem can develop. In L2 English (Caspi 2010) and L2 Finnish (Spoelman and Verspoor 2010), in writing, the complexity of certain linguistic features has been shown to develop before accuracy. The second type of interaction is a competitive relationship. This means that different subsystems may compete for the same resources, and progress in one subsystem may happen at the expense of another. For example, Verspoor *et al.* (2008) found that an advanced learner of English focused on writing alternately on vocabulary and syntactic complexity as they showed a strong trend toward a competitive relationship between Type Token Ratio and average sentence length. In the third type of interaction—a supportive relationship—the different subsystems are more coordinated and develop together, supporting each other’s growth. For example, Spoelman and Verspoor (2010) found that in writing, two complexity measures at different levels of granularity—the number of words per finite verb, a general complexity measure, and NP length, a more specific measure—correlated positively, supporting the idea that these measures were connected growers. However, as L2 development is a dynamic process, the interaction between the different subsystems may be asymmetrical,

meaning that the conditional, competitive, or supportive relationships between them may also change over time (Spoelman and Verspoor 2010).

Variability refers to changes in a variable (subsystem) within an individual over multiple measuring points (van Geert and van Dijk 2002). Thelen and Smith (1994) argued that in periods of rapid development, a subsystem might exhibit more variability than in periods of slower progress. The larger degree of variability in behavior is related to the learner's attempts to perform a task: when the learner is trying out different, old and new, ways of accomplishing the task, this results in increased variability. In line with Thelen and Smith (1994), Ellis (1994) found that in L2 learning, variability occurs more in the early stages of development than in the later stages. In the development of L2 Finnish, it has been found that case errors showed more variability in the earlier stages but that the accuracy ratios stabilized as the learner developed (Spoelman and Verspoor 2010; Tilma 2014). However, higher degrees of variability occur not only in the early stages of L1 and L2 development but also at more advanced stages, when the learner is trying out different linguistic means to convey a certain meaning (van Dijk *et al.* 2011).

According to dynamic systems theorists such as Thelen and Smith (1994), variability is a prerequisite for development and it is related to progress. Lowie *et al.* (2017) found that in L2 development, the twin with the highest degree of variability progressed more than her sister. In the multiple case study with 22 L2 English learners, there was a strong correlation between the degree of variability in holistic scores and proficiency gains (Lowie and Verspoor 2019). For progress to take place, it is necessary for the learner to try out and possibly even overuse certain linguistic features (Lowie and Verspoor 2019). When there is a regression in learner language, or when no change is taking place, variability is lower.

To summarize, DUB studies so far have shown that each learner has his or her own L2 developmental trajectory and that these trajectories are individually owned. L2 learners seem to have little in common except for some very general patterns, which may hold for most but not all learners. The question raised in this study is whether these kinds of patterns, more specifically a non-linear development of subsystems, changing interactions between subsystems, and increased variability in periods of progress, can be found in Finnish learners who try to express the concept of evaluation. In other words, the main goal of the current study is to explore whether there are similarities in the learners' trajectories when—contrary to many previous studies—the analysis is started from the meanings that the learners convey, that is, when an onomasiological approach is adopted.

THE CURRENT STUDY

Taking an onomasiological approach—that is, searching for the formal verbalizations of a given concept (Grzega 2012; Fernández-Domínguez 2019)—the current study aims to capture the changes the developing L2 system undergoes as the learner uses the language for the purposes of social interaction.

The current study will trace the development of different types of evaluative constructions in four beginner learners of Finnish.

Our research questions are as follows:

- 1 What types of constructions do the learners use to express evaluation and what types of interaction can be observed between these constructions?
- 2 What kinds of patterns of variability can be observed in the use of the different types of evaluative constructions?

Based on our earlier study (Lesonen *et al.* 2017), we hypothesize that

H1: The learners will use mainly two constructions (verbal and adjectival) to express evaluations and these constructions have a competitive relationship.

If there is a competitive relationship between the construction types, we should see that when one construction type is being explored and, therefore, used more frequently, both the token and type frequency of the other construction type will decrease. If the hypothesis about a competitive relationship is not supported, we should find a more or less equal distribution of construction types throughout the period of observation.

H2: When one construction type is being explored, this construction will show more variability compared to the other type.

If this hypothesis is supported, we should see that a construction type that is being explored and thus used more frequently shows a higher variance in token frequency than the other construction type. If the hypothesis is not supported, the variances of both types should be more or less at the same level throughout the period of observation.

METHOD

Participants

This study traces the language development of four adult Finnish L2 learners who were then attending the same language courses at a university in Finland. The background information on the participants is presented in Table 1.

To ensure that these four learners' data were comparable in terms of Finnish language proficiency at the beginning of the study, three experienced raters who were L1 speakers of Finnish evaluated the learners' first three written texts. The length of these texts ranged between 39 and 167 words (average 93 words). The criteria of the Finnish National Certificates of Language Proficiency testing system were used (the University of Jyväskylä, Center for Applied Language Studies and the Finnish National Agency for Education; scale from 1 to 6, 1 being the lowest and 6 the highest level, corresponding the levels A1–C2 in the European Framework of Reference for Languages

Table 1: Background information on the participants

Participant	Age (years)	L1	Other languages	Time of residence before the study (years)	Explicit instruction before the study
Lena	23	German	English, ^{a,b} French, ^a and Icelandic ^{a,b}	0	0
Jungo	22	Chinese (Hunanese)	Mandarin Chinese ^a and English ^a	2	One Finnish course of five ETCS, 20 hours self-studying
Alvaro	30	Spanish	English, ^a French, ^{a,b} and Russian ^a	0	0
Khadiza	31	Bangla	English, ^a Hindi, and Urdu	4	0

^aLearned in instructional setting.

^bLearned in target language-speaking community.

Table 2: Participants' L2 writing proficiency at the beginning of the study

Participant	Text 1		Text 2		Text 3		Texts 1–3 together	
	Range	Median	Range	Median	Range	Median	Range	Median
Lena	1	1	1	1	2–3	3	1–3	1
Jungo	1–2	2	1–2	2	2–3	2	1–3	2
Alvaro	1	1	1–2	1	2–3	3	1–3	1
Khadiza	1–2	2	1–2	2	2–3	2	1–3	2

(Common European Framework of Reference for Languages, Council of Europe, 2001)). For the first written text (Table 2, Text 1), the median of ratings for Khadiza and Jungo was higher than for Lena and Alvaro. However, in week 5 (Table 2, Text 3), Lena and Alvaro had caught up with Khadiza and Jungo: the median of the ratings for the third written text was higher for Lena and Alvaro than it was for Khadiza and Jungo. In other words, the initial differences in participants' language proficiency leveled out during the first five weeks of the study. The range and the median of the three ratings for the first three written texts for every participant are shown in Table 2.

All the participants took the same three Finnish courses during the study. The three courses were at the levels A1, A2, and B1 in the European

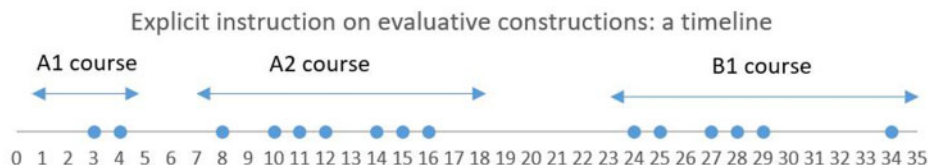


Figure 1: The timing of explicit instruction on expressing evaluation during the study.

Framework of Reference for Languages (Common European Framework of Reference for Languages, Council of Europe, 2001) (see Figure 1 for the timing of the courses). Each course consisted of 70 contact hours and additional independent work; each one was worth five ECTS (European Credit Transfer System). All three courses were taught by an L1 speaker of Finnish, the first two courses by the first author, and the third one by a colleague.

The three courses aimed to develop learners' skills in four different functions: social interaction, telling and describing, understanding and searching for information, and developing as a language learner. While evaluative constructions (that something is good or bad, desirable or non-desirable) were not the explicit focus of the course, they were taught and practised during the courses on some occasions as part of the 'Telling and describing' learning goal. The timing of the explicit instruction on these constructions is given in Figure 1. The exercises are shown in Supplementary File 1. In total, evaluative constructions were taught in 15 of the 36 weeks of the data collection. In all of these 15 weeks, except in weeks 24 and 34, both verbal and adjectival constructions were presented or used during the activities. A more detailed table showing the constructions used in each week is presented in Supplementary File 2.

Data collection

The data were collected weekly over a period of nine months, with written and spoken data alternating. The number of data points is shown in Table 3.

The data are free-response data: the participants were asked to write or talk about a certain topic. The tasks include topics like 'Write about yourself', 'Write an invitation to your birthday party', and 'What are you going to do next summer?'. All the topics are listed in Supplementary File 3. The topics were selected in accordance with course contents and the participants' language proficiency levels so they were familiar to the participants. However, they were not practised in the classroom.

The written data are handwritten. In the first half of the study, the data were collected during the contact lessons of the Finnish course and there was a time limit of approximately 20 minutes. In the second half, the data were collected in the participants' free time and there was no time limit. The length of the writing samples is on average 91 words (range: 31–152 words).

Table 3: Number of data points

Participant	Number of data points	Written data points	Spoken data points
Lena	35	17	18
Jungo	35	18	17
Alvaro	33	16	17
Khadiza	28	16	12

The spoken data were recorded in a language studio, with a recorder (Roland R-05) and once with a smartphone. The spoken data consist of both monologs and dialogs. The other person in a dialog was either another L2 speaker from the classroom or an L1 speaker of Finnish (mostly the first author of the paper, twice another Finnish instructor or a research assistant). The length of the speaking samples is on average 218 words (range: 44–518 words).

Data selection

The data were first transcribed in Word. CHAT (Codes for Human Analysis of Transcripts) format (MacWhinney 2000) was followed to the extent necessary for the analysis of this study (e.g. overlaps were not transcribed). After that, the first writer of the paper scrutinized the data for all the constructions expressing evaluation. In this study, the evaluative language is defined in line with Alba-Juez and Thompson (2014: 13) who define evaluation as

a dynamical subsystem of language, permeating all linguistic levels and involving the expression of the speaker's or writer's attitude or stance towards, viewpoint on, or feelings about the entities and propositions the s/he is talking about(.)

In line with CDST assumptions, Alba-Juez and Thompson (2014) point out that evaluative language can be seen as a subsystem of language. All constructions expressing evaluation are part of this subsystem, which can be further divided into smaller subsystems, like different construction types expressing evaluation (see section Data categorization). Linguistic levels that evaluative language permeates are the phonological, the morphological, the lexical, the syntactic, and the semantic level (Alba-Juez and Thompson 2014: 10 and 11). Alba-Juez and Thompson (2014: 10) point out that the lexical level is 'the most evident level' of evaluative language when words with 'evaluative load', such as *good* or *bad*, are used. Evaluations at this level were also most evident and frequent in the data of this study (see section Data categorization). In the corpus of this study, expressions of evaluation include expressions like:

Mut se oli kiva 'But it was nice' (attitude or stance toward/view point on entity)

Tykkään Pink Floydista ‘I like Pink Floyd’ (attitude or stance toward/view point on entity)

Mua ärsyttää kaikki ‘I’m annoyed by everything’ (feeling about entity)

He ovat tärkeä mun elämässä ‘They are important in my life’ (attitude or stance toward/view point on entity)

**Siita minua piristää* ‘That cheers me up’ (feeling about entity)

Ajattelen mun mielestä se on hyvä **idia* ‘I think it’s a good idea’ (attitude or stance toward/view point on proposition)

In the current study, interrogative expressions like *Tykkäätkö pitsasta?* ‘Do you like pizza?’ are included in the analysis. Simple yes/no statements are excluded.

In data selection, an onomasiological approach (Grzega 2012; Fernández-Domínguez 2019) was used. The term onomasiology refers to a process proceeding from notion to name, while the opposite process, semasiology, goes from name to notion (Malmkjær 1991: 291). The onomasiological approach searches for the linguistic means that are used to express a certain notion (Grzega 2012). Using the above-mentioned definition of evaluation, the first author selected all evaluative utterances based on the meaning of every utterance within its context. When using the onomasiological approach, an essential point is that the selection is based only on the meaning of the construction, not on its form. Later on, in the process of data categorization, the form of the expressions was also paid attention to (see section Data categorization). The onomasiological approach emphasizes meaning making as a central function of language (Fernández-Domínguez 2019) and by using it, we can get a better idea of the learners’ communicative needs compared to a more traditional analysis that starts with a linguistic form. Moreover, when no strict requirements for the form of the targeted construction are set, also unconventional and non-target-like learner language constructions are included in the analysis (e.g. **ei* (NEG) + adjective, **Mä tunnen että Suomi on parempi* ‘I feel that Finland is better’, or **Sitten minulla on hyvin nukkuminen* ‘Then I have well sleeping’).

Because also the unconventional, non-target-like learner language construction wanted to be included into our analysis, in the current study, Goldberg’s (2006) definition of a construction—a conventionalized pairing of form and function—has been broadened somewhat to include the L2 learners’ emergent form-meaning mappings, which might not yet seem conventional from the point of view of proficient language users. Because of these special characteristics of learner language constructions—unconventionality and impermanence—the definition of learner language construction given by Waara (2004: 53) has been adopted in this study:

A learner construction is a construction, i.e., a meaning and syntax correspondence, but which is used in a slightly unconventional manner. Although usage does not result in a communication breakdown between participants, it deviates in some way.

Examples 1 and 2 show how meaning and syntax correspondence, that is, the link between the semantic and the phonological pole might be expressed with a conventional and an unconventional construction (see Online Appendix 1 for the glossing). If the inaccuracies in the learner language constructions are small enough, the link between form and function remains clear and therefore the expressions are understandable and the communication does not break down:

- | | | | | |
|-----|--------------------|---------------|----------------|------------------|
| (1) | <i>Minä</i> | <i>pidä-n</i> | <i>enemmän</i> | <i>talve-sta</i> |
| | I | like-1SG | more | winter-ELAT |
| | I like winter more | | | |
| (2) | <i>Minä</i> | <i>pidä-n</i> | <i>enemmän</i> | <i>*talvi</i> |
| | I | like-1SG | more | *winter |
| | I like winter more | | | |

In the previous single case study (Lesonen *et al.* 2017), all of Lena's evaluative constructions in written tasks, presented in their original context, were given to a panel of three proficient speakers of Finnish, who judged whether the utterances expressed evaluation or not. Based on Lena's data, the first author scrutinized the data of the other three participants. In the case of a few problematic expressions, the Finnish-speaking panel was consulted. In total, 859 expressions of evaluation were included in the final analysis (see Table 4 for the normalized frequencies).

As pointed out by Alba-Juez and Thompson (2014), evaluation can be expressed at the lexical level by using words with 'an evaluative load' (Alba-Juez and Thompson 2014: 10). When the expressions of evaluation used by the participants of this study were explored, it turned out that almost all of them include a word with an evaluative load (e.g. *tykätä* 'like', *rakastaa* 'love', *hyvä* 'good', and *tärkeä* 'important'). In a very few expressions, the evaluative word could not be identified, like in Lena's expression *Joo, se on Suomessa ... ei lunta, ei aurinko, ei nukkuma* 'Yep, that's Finland ... no snow, no sun, no sleeping'. In this expression, (pragmatic) evaluation is covert and can be interpreted with the help of the context: in November, Lena expressed a negative attitude toward Finland because it had been so dark lately and she had been feeling tired. In this kind of expression, the speaker and the hearer have to share some knowledge; otherwise, the evaluative meaning may be lost (Alba-Juez and Thompson 2014).

Data categorization

As pointed out earlier, in virtually all evaluative expressions, the evaluative word could be identified and these expressions were categorized according to the main evaluative element of the construction, that is, the word that classes the construction as evaluative. As was the case in our earlier single-case study, it turned out that there were two main types of constructions in the data: verbal and adjectival/adverbial¹ constructions. As pointed out by Martin and White (2005: 46), verbs and adjectives express different kinds of evaluations.

Table 4: Normalized frequencies of evaluative constructions in written and spoken data: mean, standard deviation, and median

Participant	M		SD		Median	
	Written	Spoken	Written	Spoken	Written	Spoken
Lena	3.92	2.84	2.39	2.30	4.20	1.92
Jungo	4.68	3.47	3.53	2.33	3.23	2.78
Alvaro	4.66	3.22	2.45	1.36	3.90	3.29
Khadiza	4.87	5.50	2.67	3.38	4.82	4.40

Verbs describe evaluation as process, like in Example 1, adjectives instead express evaluation as quality, like in Examples 3–5. These two types of constructions are normally mutually exclusive within one utterance; in other words, for any given utterance, they are informal competition for encoding the evaluation. Because these two types of constructions express evaluation but encode it in a different way, we see them as subsystems of the evaluative language, which can be seen as a subsystem of language (Alba-Juez and Thompson 2014: 13).

In a verbal construction, the evaluative element of the construction is a verb such as *tykätä* ‘like₁’, *haluta* ‘want’, or *vihata* ‘hate’. In our analyses, the construction also encompasses all the complements of that verb (subject, object, and adverbial) and optional qualifiers (see Example 1 above, in which the verb *pitää* ‘to like₂’ forms the core of the construction).

In an adjectival construction, the adjective is the constitutive element in which the speaker expresses his or her evaluation of, or attitude toward something. Adjectives are often used in a comment clause which is a declarative sentence with a copula (VISK § 1212) and an adjective (or an adverb) used predicatively. In our analyses, the adjectival constructions also encompass the subject and the optional qualifiers (see Example 3, in which the core of the construction is the word *hyvä* ‘good’):

- (3) *ja* **kasvastudeiden o-n* *tosi* *hyvä*
 and **education (science) be-3SG* very good
 ‘and the education is very good’

In some sentence types, the adjective is used attributively (see Examples 4 and 5):

- (4) *Häne-llä* *o-n* *mielenkiintoinen* *tarina*
 He-ADE be-3SG interesting story
 ‘He has an interesting story’
- (5) *Me* *sö-i-mme* *hyvä-ä* *ruoka-a.*
 We eat-PST-3PL good-PAR food-PAR
 ‘We ate good food.’

As shown in Example 5, sometimes the verb in the adjectival construction is something other than *olla* ‘to be’. This verb is never an evaluative verb, but the evaluative or affective element of these phrases is an adjective or an adverb.

In total, 24 of the 859 expressions of evaluations could not be categorized as verbal or adjectival/adverbial constructions. In some of these expressions, a word with an evaluative load could be identified, that is, evaluation was expressed at the lexical level, like in the expression *Siellä on edistys, totta kai* ‘Of course, there is progress’. In a very few expressions, the evaluative word could not be identified, like in Lena’s evaluation on Finland presented earlier in this section. Because these expressions were not used repeatedly, they were not analyzed in more detail. These expressions are shown in [Supplementary File 4](#).

Normalizing the data

Because the texts produced by the participants varied in length, the data were normalized for text length. Both the token and type frequencies of constructions were calculated per 100 words.

Creating one spoken and written corpus

Before creating one corpus of the spoken and written data, the two types of data were compared. The descriptive statistics (mean, standard deviation, and median) are shown in [Table 4](#).

It transpired that the evaluative constructions used in the two modes were similar in frequency. Paired samples *t*-tests (Lena and Khadiza) and Wilcoxon Signed-ranks tests (Alvaro and Jungo) showed no significant differences between the spoken and written data evaluative constructions frequencies (Lena: $t(16) = 1.470$, $p = 0.161$; Khadiza: $t(11) = -0.586$, $p = 0.569$; Alvaro: $Z = -1.603$, $p = 0.109$; Jungo: $Z = -1.022$, $p = 0.307$). Because these tests require an equal number of data points in both data sets, one data point needed to be removed from each learner’s spoken or written data. The mean, standard deviation, and median reported above are based on these data sets. Means for the complete data sets before the removal of one data point each are as follows: Lena spoken data: 2.91, Jungo written data: 4.63, Alvaro spoken data: 3.26, and Khadiza written data: 5.42.

Visualizing interactions among constructions

A DUB analysis always starts with visualizing the raw data in different types of line or bar graphs to observe general trends. Then various ways may be used to detect patterns of interaction. To examine the interaction of verbal and adjectival of constructions longitudinally—as it was found out that all learners used almost exclusively verbal and adjectival constructions to express

evaluation (see section Data categorization)—the normalized frequencies were smoothed. The idea of data smoothing is to make trends in the data more clearly recognizable and this is done by decreasing the amount of variability in the data that is plotted in a graph (Gunst and Mason 1980: 39). In the current study, locally estimated scatterplot smoothing (LOESS), which is a type of local regression (Peltier 2009), was used. This method creates a LOESS curve, which is based on linear regression lines that are calculated for parts of the data by using a moving window (for linear regression, see Gunst and Mason 1908: 6–8). The data points in the center of the window have a bigger effect on the slope of the LOESS curve than the data points at the edges of the window (Harrell 2015: 29).

The smoothed trajectories reveal the interaction patterns more clearly than the raw data and they show that all learners use verbal and adjectival constructions in phases: at times verbal constructions are used more frequently than adjectival constructions, and vice versa. Based on these differences in both token and type frequencies of verbal and adjectival constructions, each learner's data were divided into verbal and adjectival phases. The smoothed data also help us to see how the frequencies of the constructions are changing over time within the phases, that is, whether the frequencies of the two types of constructions are increasing or decreasing over time.

When smoothing the data, different window sizes were used for different learners because of the differences in the number of data points. For Lena, we used a window of 12 data points ($\alpha = 0.343$), for Jungo a window of 12 data points ($\alpha = 0.387$), for Alvaro a window of 11 data points ($\alpha = 0.333$), and for Khadiza a window of 10 data points ($\alpha = 0.400$). The value for α should be set between 0.25 and 0.50 and tells how many percent of the data points fall within the used window size (e.g. for Lena 12 data points out of 35 correspond to 34.3%).

The interaction patterns found in smoothed data were again compared to the raw data, which allow us to see the interaction and the development of frequencies gleaned from the smoothed data in more detail. It is important to note that smoothed trajectories do not have a one-on-one relation with the raw data points.

Calculating and visualizing degrees of variability

We also aimed to investigate the variability patterns in the four learners' data. In this study, the variability is operationalized in terms of variance. Variance measures how much a set of numbers on average deviates from the mean. In the current study, the variance is defined as the average of the squared deviations from the mean, that is, the squared standard deviation (SD) (van Geert and van Dijk 2002). Variance is calculated for phases defined on the basis of the raw data. For each learner, the variance of different phases was compared to the variability patterns found using min–max graphs (see [Supplementary File 5](#)) and we found that they were in line with each other in virtually all

Table 5: Normalized frequencies of different constructions expressing evaluations over the whole period of observation

Participant	Total	Verbal	Adjectival	Other
Lena	118.95	58.22	57.75	2.98
Alvaro	129.90	59.39	69.53	0.98
Khadiza	136.72	69.64	59.79	7.29
Jungo	125.93	65.00	59.75	1.18

cases. This comparison ensured that the variance does not overestimate the amount of variability even though in some cases, the variance might be sensitive to the mean (van Geert and van Dijk 2002).

RESULTS

Table 5 shows the normalized frequencies of the different constructions used by each learner. As shown in Table, all four learners used mainly verbal and adjectival constructions to express evaluation and for this reason, only these two types of construction will be analyzed in detail.

First, to get an overall view of the data, we will report both the token and type frequency of the two construction types used by each learner in the course of the nine-month study. After that, we will report on each learner separately. The first learner, Lena, is discussed in great detail to show the line of reasoning (see also Lesonen *et al.* 2017). In the discussion, we will consider to what extent we may see similar developmental principles in the four learners' trajectories.

Lena

Figure 2 shows Lena's development over time of the two types of constructions. The LOESS function visualizes the changing interactions between the constructions.

The LOESS curves show clear phases in the use and interaction of the two construction types. Initially, Lena strongly prefers verbal constructions, like *Tykkään tomaatista* 'I like tomato' and *Rakastan *puhuaa *Suomea* 'I love to speak Finnish'. This phase roughly spans weeks 1–11 and can be divided into two subphases. In subphase 1a (weeks 1–6; the phases are indicated by the bar below the title of the graph), the frequencies of both constructions are increasing, but in subphase 1b (weeks 7–11), the frequencies of verbal constructions are decreasing while the frequencies of adjectival constructions, like *Jyväskylän hyvä kaupunki* 'Jyväskylä is a nice city', are increasing. In subphase 1b, these two types of constructions are, then, in a competitive relationship. After this, the token and type frequency of verbal constructions continue to

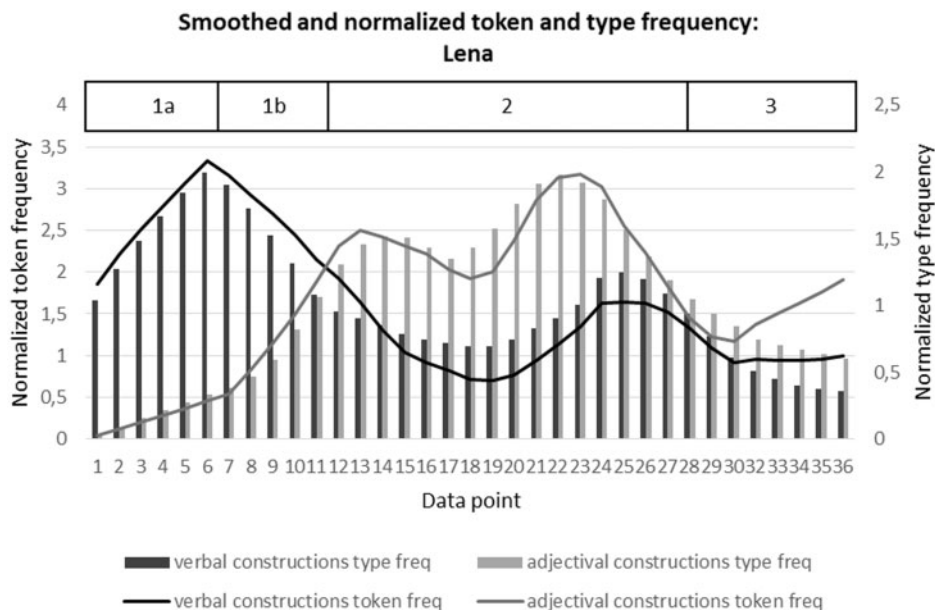


Figure 2: Lena's smoothed normalized token and type frequencies of verbal and adjectival constructions

decrease and in phase 2, around weeks 12–27, adjectival constructions are used more frequently (the token frequency is higher) and with more variability (the type frequency is higher) compared to verbal constructions. In phase 3 (weeks 28–35), the frequencies stabilize at more or less the same level and one type is not used at the expense of the other.

The smoothed trajectories in Figure 2 help us to see the general patterns of interaction between the constructions. Inspection of the raw data allows us to see the interaction and development of frequencies in more detail and to refine the phases set up on the basis of the smoothed data. Table 6 presents the refined phases and the actual numbers of type and token frequencies and the variance in token frequency. Figures 3 and 4 show the verbal and adjectival constructions used over time. It is important to note that smoothed and raw data show the development in a different way, and the data points of the smoothed trajectories do not have a one-on-one relation to the data points in the raw data figures.

Averages of the raw data (Table 6) show clearly that initially, verbal constructions are used at the expense of adjectival constructions. In subphase 1a, verbal constructions are used much more frequently than adjectival constructions: both the token and type frequency of verbal constructions are higher than those of adjectival constructions. In total six different verbal, but only one adjectival construction is used. In addition, the variance in the token frequency of verbal constructions is higher than that of adjectival constructions, indicating that on average, the token frequencies of verbal constructions

Table 6: Lena's different phases of construction use: the mean frequencies and the variance of verbal and adjectival evaluative constructions

Phase	Weeks	Token frequency ^a		Type frequency		Repertoire: how many new constructions?		Token frequency variance	
		Verb.	Adj.	Verb.	Adj.	Verb.	Adj.	Verb.	Adj.
1a	1–9	2.84	0.26	1.75	0.22	All (6)	All (1)	4.99	0.14
1b	10–12	2.41	2.16	0.95	0.91	1	3	2.30	2.39
2	13–25	1.04	2.74	0.84	1.78	4	15	0.43	2.46
3	26–36	1.19	1.32	0.73	0.79	2	5	1.63	1.80

^aThe token and type frequencies are calculated per 100 words.

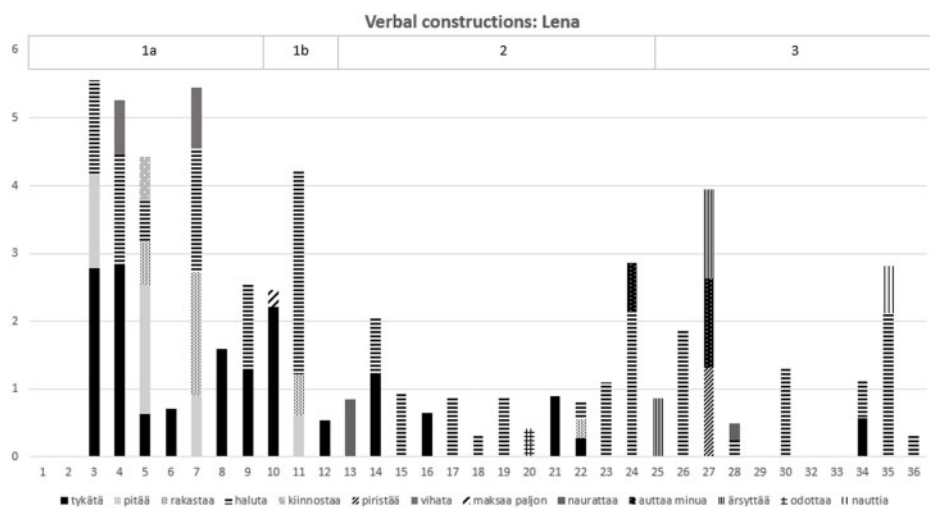


Figure 3: Different types of verbal constructions: Lena

deviate more from their mean token frequency in that period than do the token frequencies of adjectival constructions. In other words, from one session to the next, the token frequencies of verbal constructions vary on average between a wider range in that period than the token frequencies of adjectival constructions.

In subphase 1b, there seems to be a shift in Lena's language use. As the use of adjectival constructions peaks for the first time and becomes more varied, the token and type frequency of verbal constructions begins to decrease. In this phase, there is no big difference between the variance in the different

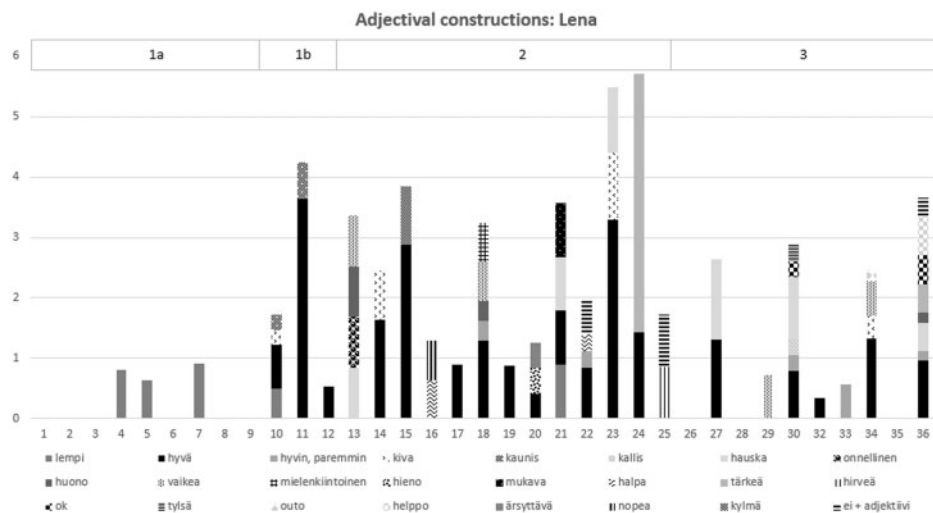


Figure 4: Different types of adjectival constructions: Lena

types of constructions: the variance in the token frequency of adjectival constructions is only slightly larger than the variance in verbal constructions.

In phase 2, both the token and type frequencies of adjectival constructions are on average higher than the token and type frequencies of verbal constructions, which are now lower than in the two earlier phases. It, therefore, seems that the more frequent and variable use of adjectival constructions happens at the cost of the use of verbal constructions. The qualitative inspection confirms this observation: at this phase, the verbal constructions that Lena prefers are the *tykätä* 'like₁' and *haluta* 'want' constructions that are familiar to her from earlier weeks (Figures 3 and 4). These two constructions are the most frequent ones in phase 2, and in weeks 14–19, they are the only verbal constructions used. Also, in this phase, the variance in the token frequencies of verbal constructions is lower than the variance in the token frequencies of adjectival constructions. In fact, the variance in the token frequencies of verbal constructions is at its lowest level.

In phase 3, the use of verbal and adjectival constructions seems to become more balanced. The average token and type frequencies of the two types do not differ much. In some texts, verbal constructions are used more frequently, and in some texts, adjectival constructions. The variance in token frequency does not differ considerably between the two construction types.

Jungo

Similar kinds of phases can also be identified in Jungo's smoothed and raw data. In Figure 5, the LOESS function visualizes the changing interactions between the constructions. Table 7 shows the averages frequencies for each

Table 7: *Jungo's* different phases of construction use: the mean frequencies and the variance of verbal and adjectival evaluative constructions

Phase	Weeks	Token frequency ^a		Type frequency		Repertoire: how many new constructions?		Token frequency variance	
		Verb.	Adj.	Verb.	Adj.	Verb.	Adj.	Verb.	Adj.
1	1–10	3.31	1.63	1.90	1.31	All (4)	All (6)	4.34	1.33
2	11–27	1.55	2.72	0.91	1.94	1	8	2.14	6.01
3	28–36	1.69	0.88	0.94	0.63	0	1	1.23	1.52

^aThe token and type frequencies are calculated per 100 words.

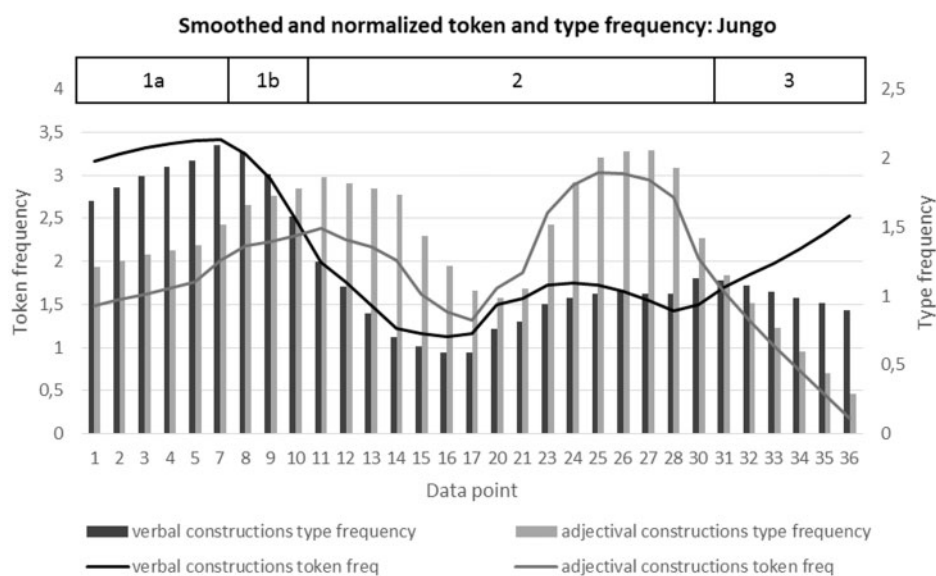


Figure 5: *Jungo's* smoothed normalized token and type frequencies of verbal and adjectival constructions

phase as established on the basis of the LOESS curves and the raw data. Figures 6 and 7 show the verbal and adjectival constructions used over time.

Also *Jungo* initially prefers verbal constructions, like *Minä rakastan Hiroa* 'I love Hiro' and *Mä myös haluaisin olla opettaja* 'I also would like to be a teacher'. This first phase roughly spans weeks 1–10 (Figure 5) and can be divided into two subphases. In subphase 1a (weeks 1–7), the frequencies of both constructions are increasing, and in subphase 1b (weeks 8–10), the frequencies of verbal constructions start to decrease. In the raw data (Table 7; Figures 6 and 7), we can also see the dominance of verbal constructions at the beginning

(higher token and type frequency and the variance for the verbal constructions in phase 1). However, the total repertoire of different constructions is bigger for the adjectival constructions.

In phase 2, approximately between weeks 11–30, adjectival constructions, like *koska se on kaunis* ‘because it’s beautiful’ and *se on ihan söpö* ‘it is quite cute’, are preferred. In this phase, the peak in the use of adjectival constructions in weeks 21–30 does not seem to affect the use of verbal constructions: the frequencies of verbal constructions stay at roughly the same level in these weeks (Figure 5). In this phase, Jungo uses almost exclusively verbal constructions that are familiar to him from earlier weeks, like *koska *minua *lakastan *minua äitiä* ‘because I love my mother’ (Figure 6).

In weeks 31–36, phase 3, there is a competitive relationship between the two types of constructions and verbal constructions are used more frequently. However, the type frequency of verbal constructions remains at a low level because Jungo relies mainly on the *haluta* ‘want’ construction.

Alvaro

Figure 8 visualizes the development of the two construction types over time. Table 8 shows the average frequencies for each phase as established on the

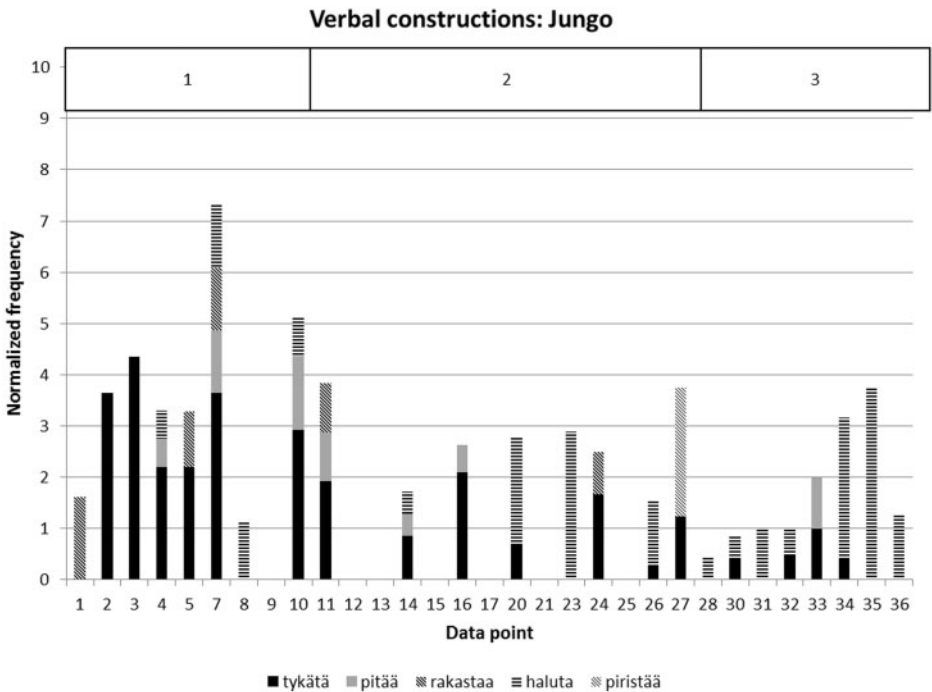


Figure 6: Different types of verbal constructions: Jungo

Table 8: Alvaro's different phases of construction use: the mean frequencies and the variance of verbal and adjectival evaluative constructions

Phase	Weeks	Token frequency ^a		Type frequency		Repertoire: how many new constructions?		Token frequency variance	
		Verb.	Adj.	Verb.	Adj.	Verb.	Adj.	Verb.	Adj.
1	1–6	3.54	1.39	2.06	1.27	All (5)	All (7)	1.41	2.47
2a	7–11	0.86	2.12	0.64	1.67	0	5	0.23	1.95
2b	13–23	1.87	1.19	1.30	0.93	0	5	1.21	0.51
3a	24–25	1.11	3.45	0.48	0.93	0	2	1.24	0.14
3b	26–36	1.30	3.19	0.90	2.09	3	10	1.71	3.37

^aThe token and type frequencies are calculated per 100 words.

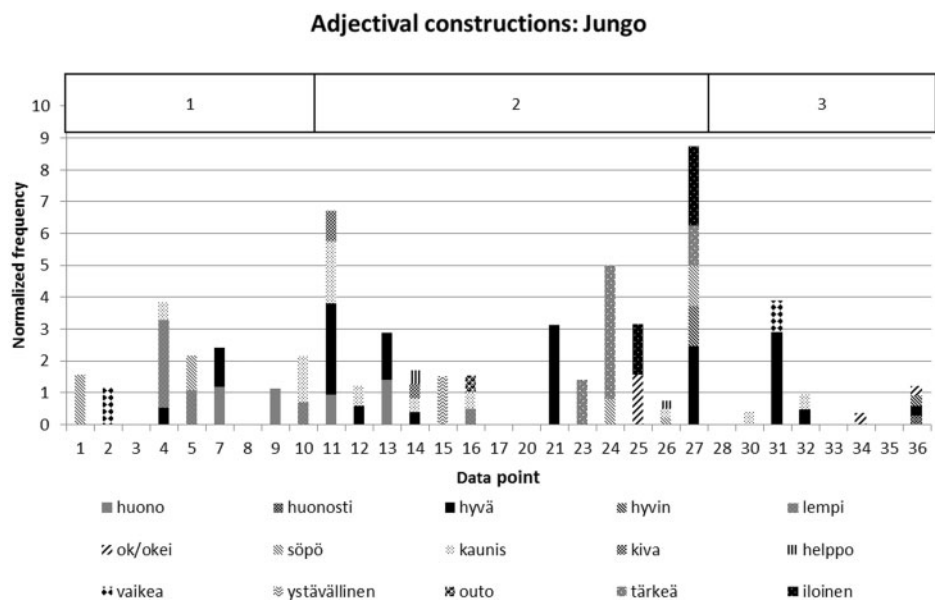


Figure 7: Different types of adjectival constructions: Jungo

basis of the LOESS curves and the raw data the two bar graphs (Figures 9 and 10) show the verbal and adjectival constructions used.

Like Lena and Jungo, Alvaro initially uses verbal constructions, like *Tykkään elokuvia* 'I like films' and *Mä pitää kauhuelokuvista* 'I like horror films' more frequently and at the expense of adjectival constructions, like

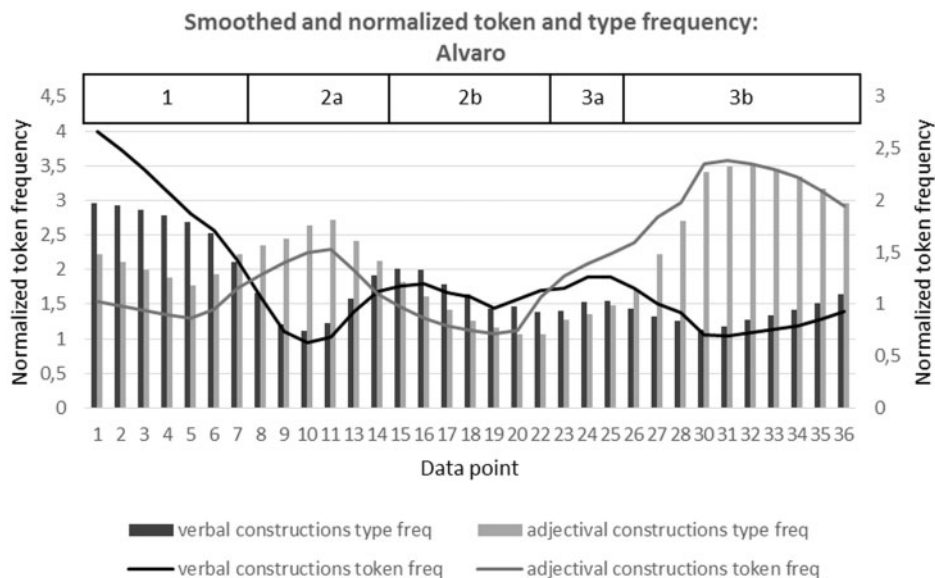


Figure 8: Alvaro's smoothed normalized token and type frequencies of verbal and adjectival constructions

hyvä idea 'a good idea' and *Puhun espanjaa *hyvä* 'I speak good Spanish'. This phase 1 roughly spans weeks 1–7 (Figure 8). The dominance of the verbal constructions is also visible in the raw data (see phase 1 in Table 8 and Figures 9 and 10). However, as with Jungo, the reliance on verbal constructions is not as strongly visible as in Lena's data. In this phase, Alvaro uses in total seven different adjectival constructions, while he uses only five different verbal constructions. Moreover, in week 5, he uses many and several different verbal and adjectival constructions (Figures 9 and 10).

In the middle of the period of observation, in phase 2, the pattern is mixed. The differences in the smoothed frequencies shown in Figure 8 are not as great as either initially or at the end of the period of observation. However, adjectival constructions are preferred at the beginning of this phase (subphase 2a). The qualitative analysis based on the raw data shows that in this phase Alvaro uses only two verbal constructions: *rakastaa* 'love' and *pitää* 'like₂', which he already used in phase 1 (Figure 9). In subphase 2b, the frequencies temporarily stabilize at more or less the same level.

In phase 3, adjectival constructions are used frequently and variably at the expense of verbal constructions. The dominance of adjectival constructions is supported by the observation that in weeks 32–36, Alvaro again relies strongly on the familiar *haluta* 'want' and *rakastaa* 'love' constructions, whereas his repertoire of adjectival constructions gets much larger (Figures 9 and 10).

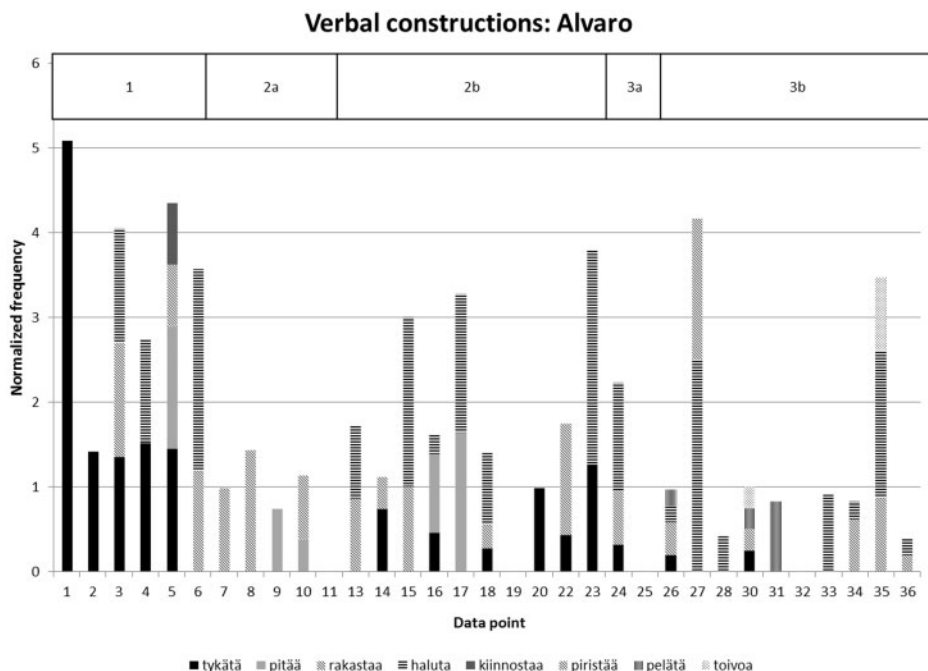


Figure 9: Different types of verbal constructions: Alvaro

Khadiza

Also for Khadiza, phases in the use of verbal and adjectival constructions can be established based on smoothed data (Figure 9) and on raw data (Table 9; Figures 12 and 13).

As shown in Figure 11, Khadiza too begins with verbal constructions, like *minä pidän *Suomi* ‘I like Finland’ and *mua ei *kinnosta *kissalle* ‘I’m not interested in cats’ when expressing evaluation. This first phase—which roughly spans weeks 1–11—can be divided into two subphases. Between weeks 1–6, in subphase 1a, the frequencies of both constructions are decreasing. Subphase 1b begins when the token and type frequencies of adjectival constructions begin to increase. This evidently has an effect on the use of verbal constructions: between weeks 7 and 11 (subphase 1b) their token frequency also increases but their type frequency decreases (see also subphase 1b in Table 9). The increase in the use of adjectival constructions, like **hyvä *saa *kesälle mutta huono *saa *talville* ‘It’s a good weather in the summer but bad in the winter’ thus creates a competitive relationship regarding the diversity of adjectival and verbal constructions (i.e. the type frequency). In this phase, Khadiza relies strongly on the *pitää* ‘like₂’ and *tykätä* ‘like₁’ constructions, which are familiar to her from earlier weeks, and she uses only one new verbal construction.

Table 9: Khadiza’s different phases of construction use: the mean frequencies and the variance of verbal and adjectival evaluative constructions

Phase	Weeks	Token frequency ^a		Type frequency		Repertoire: how many new constructions?		Token variance	
		Verb.	Adj.	Verb.	Adj.	Verb.	Adj.	Verb.	Adj.
1a	1–8	3.19	0.49	1.98	0.49	All (7)	All (5)	2.47	0.58
1b	10–14	5.01	4.66	1.27	1.94	1	4	0.44	1.36
2a	15–18	0.10	1.93	0.10	1.63	0	3	0.02	0.07
2b	20–23	5.13	0.60	1.79	0.60	0	1	2.52	0.06
3	24–34	0.96	4.18	0.44	2.21	1	12	1.34	5.56
4	35	10.06	1.52	1.52	1.52	0	1	n.a.	n.a.

^aThe token and type frequencies are calculated per 100 words.

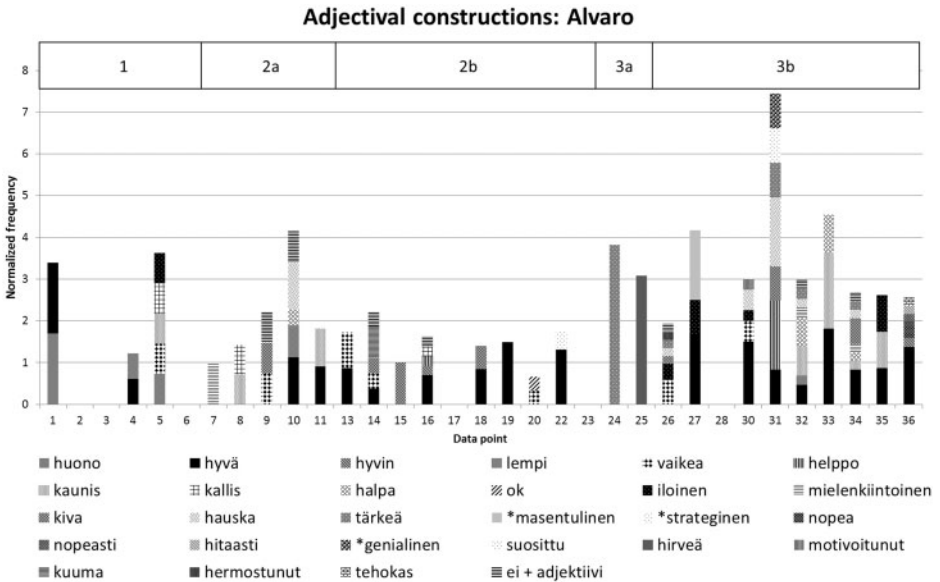


Figure 10: Different types of adjectival constructions: Alvaro

In the middle of the period of observation, namely in phase 2, the smoothed trajectories do not show a competitive relationship between the construction types, and one type is not used at the expense of the other. This second phase roughly spans weeks 14–23 (see Table 9 for a finer-grained division of the phases).

Toward the end of the period of observation, Khadiza has a clear adjectival phase in her use of constructions. In this phase of reduced use of verbal

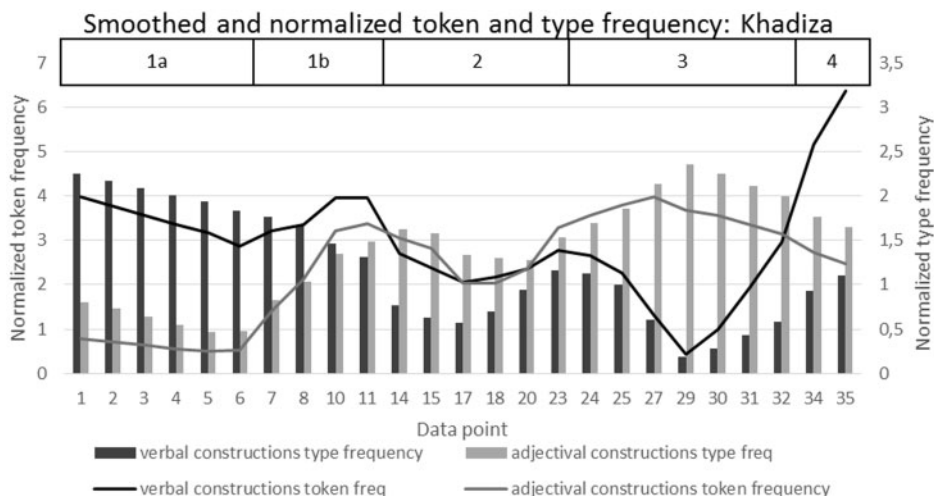


Figure 11: Khadiza's smoothed normalized token and type frequencies of verbal and adjectival constructions

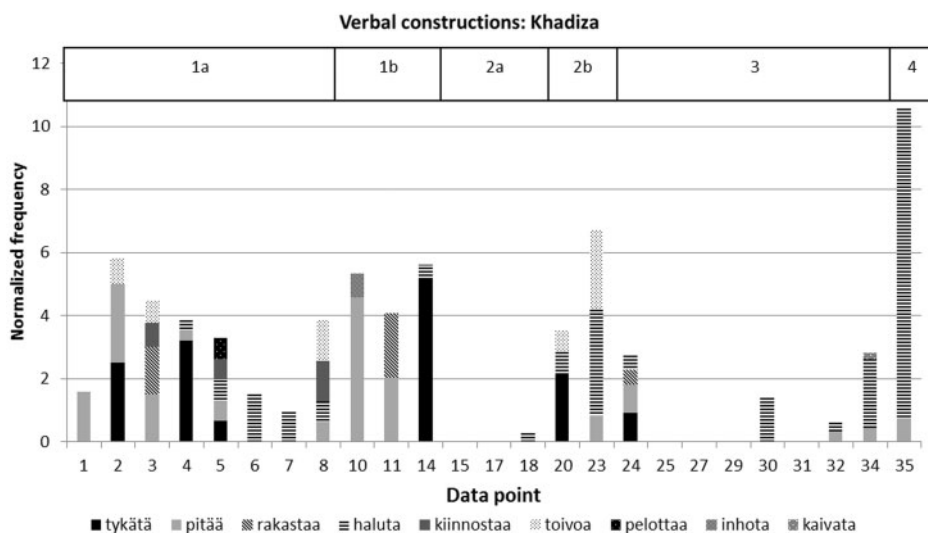


Figure 12: Different types of verbal constructions: Khadiza

constructions, we can again observe a more frequent reuse of certain verbal constructions: the entrenched *haluta* 'want' and *pitää* 'like₂' constructions are used frequently (Figure 12). The relatively sharp increase in the token frequency of verbal constructions from week 29 on, showed by the smoothed frequency, is due to a peak in the verbal *haluta* 'want' construction in week 35, which is regarded here as its own phase.

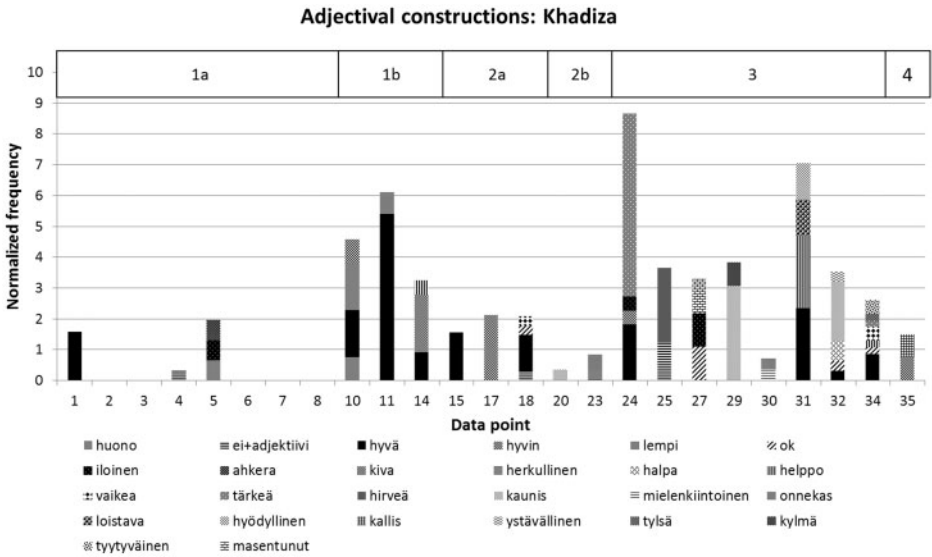


Figure 13: Different types of adjectival constructions: Khadiza

DISCUSSION

Taking a DUB approach, we followed four learners of Finnish as an L2 longitudinally and aimed to discover whether any similar patterns of interaction and variability could be found in these learners’ developmental trajectories. We expected to find highly individual learning paths with some generalizable principles, namely a non-linear development of subsystems, a competitive interaction between subsystems, and increased variability in periods of progress. In this section, the two initial hypotheses will be dealt with separately.

H1: the learners will use mainly two constructions (verbal and adjectival) to express evaluations and these constructions have a competitive relationship

This hypothesis is supported. All four learners used verbal and adjectival constructions almost exclusively to express evaluation. At the level of the utterance, these constructions are normally mutually exclusive, meaning that when an evaluative verb, such as *tykätä* ‘like’, is used, an adjective, such as *hyvä* ‘good’, is not used, and the other way around. Our study showed that also at the level of a text, all Finnish L2 beginners used these two types of constructions alternatively in certain phases of development.

For all learners, we found several phases in which verbal and adjectival constructions compete with each other or one type is used at the expense of the other. In the first weeks, all learners use verbal constructions at the expense of adjectival constructions. After this, the learners show some differences in their trajectories. In Lena’s and Jungo’s data, an adjectival phase follows

immediately after the verbal phase, in Alvaro's and Khadiza's developmental paths, the adjectival phase is only visible toward the end of the observation period. A significant feature in all four learners' adjectival phases is that when adjectival constructions are used more than before, verbal constructions are used either less frequently or in a less diverse way compared to their use in earlier phases.

These findings are in line with CDST findings in other longitudinal studies showing that when one aspect of a linguistic system is developing, there might be a regression in another aspect. It has been shown, for example, that for an advanced learner of English, vocabulary and syntactic complexity competed with each other: progress in one led to regress in others (Verspoor *et al.* 2008).

The fact that one construction is focused on at the expense of another would be expected within a CDST approach as the learner has limited attentional resources. This assumption is not new to Second Language Development (SLD) theory as it has been frequently referred to in cognitive approaches to SLD. For example, Skehan and Foster (1997) argued that learners are limited in their ability to coordinate and control attentional resources and can therefore not attend to different aspects at the same time, thus having to prioritize one language subsystem over others. This might be especially clearly visible in the language use of beginning learners (Skehan and Foster 1997). However, one essential difference between cognitive approaches to SLD and CDST is that in CDST the goal is not to explain causes or predict interaction but rather to describe the actual process.

H2: When one construction type is explored, this construction will show more variability compared to the other type

This hypothesis is partly supported. For Lena, Jungo, and Khadiza, in the verbal phase, the token frequency variance of verbal constructions is higher than that of adjectival constructions. In the adjectival phase, the token frequency variance of adjectival constructions is higher. This shows that there is more variability from text to text in the frequency of the construction type that is being explored. For Alvaro, the pattern of token frequency variance is not as clear as for the other learners. In his initial verbal phase, adjectival constructions exhibit more variability. However, in his adjectival phase in the last weeks, the token frequency variance of adjectival constructions is higher than that of verbal constructions, which is in line with our hypothesis.

This finding confirms the commonly held hypothesis in the development of dynamic systems, that a system that is in the phase of rapid progress exhibits more variability than a system that is not progressing as rapidly in that period (Ellis 1994; Thelen and Smith 1994; Verspoor *et al.* 2008; van Dijk *et al.* 2011). Exploring and trying out the new ways to express something lead to instability of the system and consequently to a growth in variability. Therefore, in CDST, variability is seen as a sign of development and is related to progress. This kind of variability pattern has been empirically established in L2 development. For example, a strong correlation between variability in holistic scores and

proficiency gains has been reported (Lowie and Verspoor 2019) and it has been found that a twin who showed more variability made more progress (Chan *et al.* 2015).

The non-linear patterns with variability patterns found in our study are not new to SLD studies either as variability was studied especially in the 1980s. In these studies—inspired by variationist research in sociolinguistics (Labov 1963)—the focus was mainly on explaining the causes of this variability, such as the task conditions (see Tarone (1983) for a review). This is not the aim of CDST studies. In CDST, patterns of variability are assumed to give insight into development. Ellis (1994), one of the first to note this in SLD, found that there was some degree of free variability, variability that could not be attributed to any known linguistic, situational, or psychological factor. In sum, variability is an intrinsic property of a developing system: free exploration of performance generates variability and it can be seen as a prerequisite for and a sign of development.

The current study contributes to our understanding of L2 as a complex dynamic system by showing that competitive interaction of subsystems and variability within these subsystems are found in the L2 also when a subsystem is viewed from the point of view of meaning and not form, that is when the onomasiological approach is used. In previous CDST-oriented studies, accuracy and complexity have been seen as subsystems (Caspi 2010; Spoelman and Verspoor 2010; Tilma 2014; van Geert and Verspoor 2015). This study suggests that the two developmental characteristics, competition and variability, may be considered as general features of a developing L2 because they were found when different, although nested, subsystems than in previous studies were studied.

In sum, this study was able to show that there are some common characteristics in the four learners' development despite the learners' clearly individually owned trajectories and other individual factors such as L1 or length of stay in Finland. However, in this study, the participants learned Finnish in the same instructional context, which may have influenced the fact that all of the learners initially used more verbal constructions. At the beginning of the study, the classroom activities that focused on expressing evaluation presented some adjectival constructions but emphasized verbal constructions (Exercises in weeks 3 and 4: how can you express liking and disliking in Finnish? Find out what other students like. What kind of music do you like/do not you like?). Therefore, the more frequent use initially of verbal constructions at the expense of adjectival constructions could partly be explained as instructional effects. However, later on in the study, both verbal and adjectival constructions were presented and used in the classroom activities (see [Supplementary File 2](#)). Since after the first course, there are differences between learners in the timing of the adjectival phase, it could be argued that later on the instruction did not play a big role in the emergence of different phases.

The overall conclusion is that despite the individual learning trajectories, these four beginning learners of L2 Finnish showed some similar patterns in their development. Our findings, therefore, confirm the assumption put forward by Larsen-Freeman and Cameron (2008: 10) that ‘each individual achieves the success that he or she does in a unique way’. However, the very general characteristics of L2 development found in the current study—namely the non-linearity of development and the increased variability when progress is taking place—might be something that could be generalized to a bigger group of learners. This might be the case even though, broadly speaking, findings from individual case studies cannot be generalized to groups (and vice versa). Because ‘no two situations can be similar enough to produce the same behavior’ (Larsen-Freeman and Cameron 2008: 16), we cannot predict the course of L2 development, but the value of case studies is in their potential to describe and hence help us to understand the process of L2 development (Lowie and Verspoor 2019).

Limitations

When analyzing this kind of data, task effects will play a role. There are texts in which the given task certainly affected the constructions that the learners used. For example, the task in week 24, ‘What is important in your life?’ inevitably triggered the *tärkeä* ‘important’ construction in every learner’s speech. In week 14, Khadiza had a speaking partner who extensively used the *tykätä* ‘like₁’ construction, which possibly played a role in the high frequency of that construction in Khadiza’s own speech. However, since all of the participants had the same tasks and there are inter-individual differences in the frequencies of constructions used in any given week, we may conclude that the patterns found in the learners’ developmental paths are not primarily due to the tasks. Besides, the patterns are also visible in the smoothed trajectories, where any task effects would be filtered out due to the smoothing.

One apparent limitation of case studies is the generalizability of the results to other learners: by definition, individual learning trajectories cannot be generalized to other learners (Lowie and Verspoor 2019). However, there are some patterns that seem to be generalizable in L2 development. The findings of this study are in line with some earlier CDST-inspired studies in that they show the non-linearity of development (alternation of progress and regress), increased variability when progress is taking place, and competitive relationships between different constructions.

CONCLUSION

The aim of this study was to investigate the development and interaction of formal verbalizations of evaluation used by four beginning L2 learners of Finnish.

We may conclude that our onomasiological approach has been able to confirm that competition may occur among constructions and degrees of variability are signs of development despite the fact that each learner has his or her individual trajectory. The study showed that at certain points of development, different evaluative linguistic constructions are in a competitive relationship with each other, or one type is used at the expense of the other. These findings, therefore, support the view of a developing L2 as a dynamic system in which a change in one component affects the other interconnected components of the system (Verspoor and Behrens 2011). The second major finding of this study was that during phases of intensive progress, more variability occurred in the frequencies of the constructions. This finding is in line with earlier findings of variability patterns in a developing L2 (Spoelman and Verspoor 2010; Tilma 2014; Lowie and Verspoor 2019) and adds to the growing body of research that indicates greater variability in phases of rapid progress.

SUPPLEMENTARY DATA

Supplementary material is available at *Applied Linguistics* online.

NOTE

1 For the clarity of presentation, the group of adjectival/adverbial

constructions will be referred to as adjectival constructions later on in the text.

FUNDING

This study has been funded by the Faculty of Humanities of the University of Jyväskylä, Ellen and Artturi Nyyssönen Foundation, Ella and Georg Ehrnrooth Foundation, and the Finnish Cultural Foundation: Central Finland Regional Fund.

REFERENCES

- Alba-Juez, L. and Thompson G. 2014. 'The many faces and phases of evaluation' in G. Thompson and L. Alba-Juez (eds): *Evaluation in Context*. John Benjamins Publishing Company.
- de Bot, K. and Larsen-Freeman D. 2011. 'Researching second language development from a dynamic systems theory perspective' in M. Verspoor, K. de Bot, and W. Lowie (eds): *A Dynamic Approach to Second Language Development: Methods and Techniques*. John Benjamins Publishing Company.
- de Bot, K., Verspoor M., and Lowie W. 2005. 'Dynamic systems theory and applied linguistics: The ultimate "so what"?' *International Journal of Applied Linguistics* 15/1: 116–18.
- Caspi, T. 2010. 'A dynamic perspective on second language development', Doctoral thesis, University of Groningen.
- CEFR. 2001. *Common European Framework for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Chan, H., Verspoor M., and Vahtrick L. 2015. 'Dynamic development in speaking versus writing in identical twins,' *Language Learning* 65/2: 298–325.

- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford University Press.
- Eskildsen, S. W. 2009. 'Constructing another language—usage-based linguistics in second language acquisition,' *Applied Linguistics* 30/3: 335–57.
- Fernández-Domínguez, J. 2019. 'The onomasiological approach' in M. Aronoff (ed.): *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.
- Goldberg, A. E. 2006. *Constructions at Work - The Nature of Generalization in Language*. Oxford University Press.
- Grzega, J. 2012. 'Lexical - semantic variables' in J. M. Hernández-Campoy and J. C. Conde-Silvestre (eds): *The Handbook of Historical Sociolinguistics*. Blackwell Publishing Ltd.
- Gunst, R. and Mason R. 1980. *Regression Analysis and Its Application*. Marcel Dekker, Inc.
- Harrell, F. 2015. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer.
- Housen, A., Kuiken F., and Vedder I. 2012. 'Complexity, accuracy and fluency' in A. Housen, F. Kuiken, and I. Vedder (eds): *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. John Benjamins Publishing Company.
- Labov, W. 1963. 'The social motivation of a sound change,' *Word* 19/3: 273–309.
- Langacker, R. W. 2009. 'A dynamic view of usage and language acquisition,' *Cognitive Linguistics* 20/3: 627–40.
- Larsen-Freeman, D. 2013. 'Complexity theory/dynamic systems theory' in P. Robinson (ed.): *The Routledge Encyclopedia of Second Language Acquisition*. Routledge/Taylor & Francis Group.
- Larsen-Freeman, D. and Cameron L. 2008. *Complex Systems and Applied Linguistics*. Oxford University Press.
- Lesonen, S., Suni M., Steinkrauss R., and Verspoor M. 2017. 'From conceptualization to constructions in Finnish as an L2: A case study,' *Pragmatics & Cognition* 24/2: 212–62.
- Lowie, W., van Dijk M., Chan H., and Verspoor M. 2017. 'Finding the key to successful L2 learning in groups and individuals,' *Studies in Second Language Learning and Teaching* 7/1: 127–48.
- Lowie, W. and Verspoor M. 2015. 'Variability and variation in second language acquisition orders: A dynamic reevaluation,' *Language Learning* 65/1: 63–88.
- Lowie, W. and Verspoor M. 2019. 'Individual differences and the ergodicity problem,' *Language Learning* 69/s1: 184–206.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd edn. Lawrence Erlbaum Associates.
- Martin, J. R. and White P. R. R. 2005. *The Language of Evaluation. Appraisal in English*. Palgrave Macmillan.
- Peltier, J. 2009. 'LOESS smoothing utility,' PTS (Computer Software), available at peltiertech.com/loess-utility-awesome-update/. Accessed 26 April 2018.
- Skehan, P. and Foster P. 1997. 'Task type and task processing conditions as influences on foreign language performance,' *Language Teaching Research* 1/3: 185–211.
- Spoelman, M. and Verspoor M. 2010. 'Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish,' *Applied Linguistics* 31/4: 532–53.
- Tarone, E. 1983. 'On the variability of interlanguage systems,' *Applied Linguistics* 4/2: 142–63.
- Thelen, E. and Smith L. B. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press.
- Tilma, C. 2014. 'The dynamics of foreign versus second language development in Finnish writing,' Doctoral thesis, University of Jyväskylä.
- University of Jyväskylä. (n.d.). 'Yleiset kielitutkinnot. Kirjoittamisen arviointikriteerit,' available at <https://www.jyu.fi/hytk/fi/laitokset/solki/yki/yleista/tietoakielitutkinnoista/arviointi>.
- van Dijk, M., Verspoor M., and Lowie W. 2011. 'Variability in DST' in M. Verspoor, K. de Bot, and W. Lowie (eds): *A Dynamic Approach to Second Language Development: Methods and Techniques*. John Benjamins Publishing Company.
- van Geert, P. 2008. 'The dynamic systems approach in the study of L1 and L2 acquisition: An introduction,' *The Modern Language Journal* 92/2: 179–99.
- van Geert, P. and van Dijk M. 2002. 'Focus on variability: New tools to study intra-individual variability in developmental data,' *Infant Behavior and Development* 25/4: 340–74.
- van Geert, P. and Verspoor M. 2015. 'Dynamic systems and language development' in B.

- MacWhinney and W. O'Grady (eds): *The Handbook of Language Emergence*. Wiley Blackwell.
- Verspoor, M. and Behrens H. 2011. 'Dynamic systems theory and a usage-based approach to second language development' in M. Verspoor, K. de Bot, and W. Lowie (eds): *A Dynamic Approach to Second Language Development. Methods and Techniques*. John Benjamins Publishing Company.
- Verspoor, M., Lowie W., and van Dijk M. 2008. 'Variability in second language development from a dynamic systems perspective,' *The Modern Language Journal* 92/2: 214–31.
- Hakulinen, A., Vilkkumäki M., Korhonen R., Koivisto V., Heinonen T. R., and Alho I. 2004. *Iso suomen kielioppi*. Comprehensive Finnish Grammar Online. Suomalaisen Kirjallisuuden Seura. VISK, Electronic Database, available at <http://scripta.kotus.fi/visk>. Accessed 29 December 2017.
- Waara, R. 2004. 'Construal, convention, and constructions in L2 speech' in S. Niemeier and M. Achard (eds): *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*. De Gruyter Mouton, pp. 51–75. <http://scripta.kotus.fi/visk/etusivu.php?s=viittaaminen>.