

Epistemic constructions in L2 Norwegian: a usage-based longitudinal study of formulaic and productive patterns*

PAULINA HORBOWICZ

Adam Mickiewicz University, Poznań, Poland and Inland Norway University of Applied Sciences, Norway

AND

MARTE NORDANGER

Inland Norway University of Applied Sciences, Norway

(Received 4 December 2020 – Revised 6 May 2021 – Accepted 6 May 2021 –
First published online 10 June 2021)

ABSTRACT

This paper addresses the development of epistemic verb–argument constructions in L2 Norwegian in four learners from a usage-based perspective. Usage-based theories hold that language learning is a gradual process of schematization. Recent research has pointed out that adult L2 learning may start out from both lexically specific and productive patterns, but also that formulaic language and semi-fixed patterns can persist for a long time in an L2. The aim of the present study is to trace how the schematization process unfolds in dense longitudinal data collected from learners in their second semester of intense Norwegian language studies, and to explore the interaction between formulaic and productive patterns in this period of language learning. The analyses show that the learners in general employ a limited repertoire of epistemic verbs, mainly *tro* ‘think’ and *vite* ‘know’. The level of productivity of *tro* and *vite* constructions varies across the learners: while one learner shows increasingly productive use of constructions with both verbs, other learners rely on semi-fixed construction patterns. A general conclusion is thus that formulaic and semi-fixed patterns are not restricted to initial phases of L2 learning and should be an object of attention at all levels of L2 competence.

[*] The authors have contributed equally. Address for correspondence: Department of Scandinavian Studies, Faculty of Modern Languages and Literatures, Adam Mickiewicz University, Collegium Novum, al. Niepodległości 4, 61-874 Poznań, Poland.
paulina.horbowicz@inn.no

KEYWORDS: usage-based learning trajectory, L2 Norwegian, epistemic verb–argument constructions, formulaic language, productivity

1. Introduction

Usage-based approaches to language and language learning hold that linguistic knowledge is the result of “the cognitive organization of one’s experience with language” (Bybee, 2006, p. 711), and that learning emerges from usage events (Langacker, 1987). Researchers have suggested that second language development, like first language development (Tomasello, 2003; Dabrowska & Lieven, 2005), starts out from lexically specific utterances that gradually schematize (Ellis, 2002; Eskildsen, 2009, 2012, 2015), rather than being productive and abstract from the start. Still, the L2 research is not fully consistent in this respect, and several studies have pointed out that some constructions are used with a high degree of productivity from early on (Roehr-Brackin, 2014; Eskildsen, 2015; Lesonen, Steinkrauss, Suni, & Verspoor, 2020a), whereas others develop from fixed and formulaic expressions. On the other hand, since learners do not always move beyond semi-fixed patterns, Eskildsen (2009) has proposed that full schematization may not be the endpoint of L2 development.¹ The aim of the present study is to pursue the schematization process of intermediate learners on their way to developing a more advanced L2 competence, and subsequently to explore the role of fixed and formulaic language in the post-initial phases of L2 development.

To meet these aims, the present paper reports on a longitudinal study comparing the development of form–meaning mappings in the epistemic constructions of the spoken Norwegian of four adult learners during their second semester of Norwegian language studies.² Norwegian is an understudied language in L2 development, and the present study thus provides new cross-linguistic data to the field. We approach epistemicity from a constructionist perspective (e.g., Goldberg, 1995, 2006; Bybee, 2006; Ellis & Ferreira-Junior, 2009a, 2009b) by focusing on grammatical constructions known as epistemic verb–argument constructions in which a group of epistemic verbs, such as *tro*, *vite*, and *anta* (‘think’, ‘know’, and ‘suppose’), occur.³ The primary function of epistemic expressions is to signal the truth-value of a

[1] Recent research into L1 acquisition even questions whether mature grammars operate on abstract patterns at all (Ambridge, 2020).

[2] The data were collected in cooperation with Gunhild Tveit Randen (Inland Norway University of Applied Sciences) as part of a larger project addressing construction development in L2 Norwegian.

[3] We apply the term ‘verb–argument construction’ descriptively without claiming that epistemic verb constructions form one unified group of constructions or can be assumed to represent one distinct schematic pattern similar to that of, for instance, the ditransitive construction (VOO) (cf. Goldberg, 2006).

given proposition (Lyons, 1977), but they also serve important discourse regulating and co-operative functions in interaction. Ochs (1996) argued that epistemic constructions serve as one of “the basic linguistic resources for constructing/realizing social acts and social identities” (p. 420). As such, epistemic constructions play an important part in reaching one’s interactional goals, performing agency, and negotiating social identity. Against this backdrop, we consider developing an epistemic repertoire a vital aspect of mastery in an L2, which deserves attention from educational, linguistic, and acquisitional points of view. In the present paper, we aim to explore the learners’ total repertoire of epistemic verb–argument constructions before tracing the developmental process of the most frequently occurring form–meaning mappings, that is, constructions with *tro* ‘think’ and *vite* ‘know’, in order to look more closely into the schematization process of each specific construction.

2. Background

This section reviews previous research informing the present study: the usage-based learning trajectory (Section 2.1) and epistemic verb constructions in Norwegian (Section 2.2). We propose two research questions targeting central features of the L2 developmental process across time and learners (Section 2.3).

2.1. THE USAGE-BASED LEARNING TRAJECTORY

In usage-based linguistics, constructions are seen as the basic linguistic unit, and language is considered “a structured inventory of constructions as conventionalized form–meaning pairings used for communicative purposes” (Ellis & Ferreira-Junior, 2009b, p. 188). Constructions may vary in their degree of both complexity and abstractness (e.g., Nistov, Gustafsson, & Cadierno, 2018; Diessel, 2019). Researchers have suggested that construction learning evolves from lexically specific utterances, that is fully fixed expressions such as *Jeg vet ikke* ‘I don’t know’ through semi-fixed patterns like *Jeg vet ikke X* ‘I don’t know X’, to fully abstract and schematic patterns, such as SUBJ + VERB + OBJ (e.g., Tomasello, 2003; Ellis & Ferreira-Junior, 2009a). This path of learning has been referred to as ‘the usage-based learning trajectory’ (see Roehr-Brackin, 2014; Lesonen et al., 2020a).

Fixed expressions, also called formulas, are recurring fixed sequences of words used with similar communicative functions, which may occur as an independent utterance or as a basis for a specific instantiation of a construction (Eskildsen, 2009). Formulas arise in the process of ‘entrenchment’ in which structures are strengthened in memory as a result of use (Langacker, 1987, p. 50): the more frequently a sequence or a structure is used, the higher the degree of entrenchment. Conversely, schematicity refers to the level of specificity and abstractness of a given construction: the less specific and less

restrictive, the more schematic. As Langacker (1987) noted: “A schema is thus abstract [...] in the sense of providing less information and being compatible with a broader range of options, even should it cover the same domains and basic properties as its elaborations” (p. 132). A construction allowing several verbs and a variety of complements, as exemplified above, would thus be more schematic than a fully lexically specific expression. Schematization and productivity are dependent on recurring meetings with similar exemplars of a construction and the subsequent gradual abstraction and generalization among similar types (Bybee, 2006; Ellis, 2006). In mature grammar, constructions are stored at different levels of abstractness, which means that the same utterance can be retrieved as a lexically specific utterance or generated from an abstract schema (cf. Langacker, 2000; Dabrowska & Lieven, 2005). In Eskildsen’s words, this multi-layeredness allows for “the cohabitation in grammar of abstract schematized representations and their concrete instantiations” (2020, p. 108).

The schematization of verb-centred constructions has been the focus of research since the mid-1990s. There exists a growing body of research on the development of verb–argument constructions (VACs) in both L1 (Tomasello, 1992; Ninio, 1999; Goldberg, Casenhiser, & Sethuraman, 2004; Lieven & Tomasello, 2008) and L2 learning (Ellis & Ferreira-Junior, 2009a, 2009b; Römer & Berger, 2019; Römer, Skalicky, & Ellis, 2020). According to Goldberg (1995), VACs form basic sentence type constructions and encode “event types that are basic to human experience”, such as “someone causing something, something moving, someone possessing something” (p. 37). The famous sentence *Tom sneezed the napkin off the table* – a specific instantiation of the abstract construction X causes Y to move Z (verb–object–locative; Ellis & Ferreira-Junior, 2009a, p. 371) – reflects how the construction as a whole, not only the main verb of a sentence, is a carrier of meaning. Rather, the verb is provided with its meaning by the construction in which it occurs.

Tomasello (1992, 2000) argued that children’s early grammars are based on specific verb constructions developing independently of other verbs, referred to as ‘verb islands’. Research into the development of VACs in L1 and L2 learning has documented that different types of VACs often develop from so-called path-breaking verbs, which are frequent, prototypical, and generic in meaning (Ninio, 1999; Goldberg et al., 2004; Ellis & Ferreira-Junior, 2009a). Goldberg (2006) has further suggested that construction learning is optimized by a Zipfian type/token frequency of verb distribution that provides the learner with one very frequent exemplar of each abstract construction schema. Additionally, these highly frequent verbs also account for the “lion’s share of tokens in each of the constructions considered” in children’s early VAC production (Goldberg et al., 2004, p. 307). This insight was replicated for L2 learning by Ellis and Ferreira-Junior (2009a), who found that for verb–locative constructions, *go* was by far more frequent than other verbs, whereas *put* and *give* largely

outnumbered the use of other verbs in the verb–object–locative construction (VOL) and in the ditransitive construction (VOO), respectively. In line with Ninio (1999), Ellis and Ferreira-Junior (2009a) concluded that the most frequent verbs were generic in meaning, but they were not the most prototypical verbs occurring in each VAC. When extending the analyses to include the entire VAC (subject–verb–argument), a set of prototypical sequences emerged, argued to represent early uses of L2 English VACs, such as *I went to the shop* and *put it in the bag* (Ellis & Ferreira-Junior, 2009b, p. 211). Similar patterns have been detected in recent research. In a large-scale corpus study targeting the development of 19 different VAC-constructions, such as *V in n* and *V with n*, in written L2 English in two learner groups (L1 German and L1 Spanish) across proficiency levels, Römer and Berger (2019) found an early reliance on verbs with generic and prototypical meaning in both L1 groups. Their results further confirmed the importance of pathbreaking verbs (particularly for *V in n*). At more advanced stages, however, the use of VACs became more varied, more similar to ‘native’ use, and the differences between L1 groups grew.

Ellis and Ferreira-Junior (2009a, 2009b) and Römer and Berger (2019) primarily focused on frequency and typical VAC associations across learners, rather than on individual developmental trajectories. Studies within a usage-based approach following individual learners over time have demonstrated that L2 learners’ construction development may indeed follow the usage-based learning trajectory, be initially lexically specific (Eskildsen, 2009, 2012, 2014, 2015; Roehr-Brackin, 2014; Lesonen, Suni, Steinkrauss, & Verspoor, 2018), and develop on the basis of a few exemplars in the ambient language (e.g., Ellis & Ferreira-Junior, 2009b). The prevalence of formulaic and semi-fixed language has further been documented in studies of L2 negated constructions (Eskildsen, 2012) and *wh*-questions (Eskildsen, 2015). Eskildsen (2009) showed that the use of *can* constructions developed from the multiword expression *I can write*, while other verbs were included in the construction frame only later. Based on this study, Eskildsen (2009) argued that semi-fixed constructions are prominent in the L2 repertoire, a conclusion that has been further supported by subsequent studies (Eskildsen, 2014, 2017, 2020) applying the trace-back methodology originally developed for L1 acquisition (cf. Dabrowska & Lieven, 2005; Lieven, Salamo, & Tomasello, 2009). Eskildsen (2014, 2017) showcases how a learner’s construction repertoire in spoken longitudinal data both at beginning and more advanced stages could be traced back to earlier utterances by the application of simple operations such as substitution of construction parts. These developmental processes are further strengthened in analyses of another learner’s L2 repertoire in Eskildsen (2020). As such, L2 development is shown to be a gradual and stepwise process where novel utterances may be accounted for as expansions of already existing

patterns with one recycled and one substituted part. Semi-fixed constructions, i.e., constructions consisting of fixed and open slots, are key to this process (cf. Eskildsen, 2009).

In a previous study (Horbowicz, Nordanger, & Randen, 2020), we followed two adult learners, Linda and Emilio, through their second semester of Norwegian studies. Although epistemic verb–argument constructions with *tro* and *vite* appeared to be productive in terms of possible complements in the learners' spoken language, these constructions seemed to have a fixed base in the L2 of one of the learners. These basic sequences consisted of *jeg tror X* 'I think X' or *jeg tror det er X* 'I think it is X' and *jeg vet ikke X* 'I don't know X'. This finding corresponds well with patterns found in research on L1 English, as well as in studies on the L2 Swedish phrase *det är* 'it/there is', suggesting that this sequence is learned and automatized as a formulaic chunk (Bartning & Hammarberg, 2007; Hammarberg, 2008). Diessel and Tomasello (2001) have also found formulaicity in children's early uses of *think* and *know*, consisting of little variation in the subject slot (predominantly the first person pronoun *I*) or an entirely missing subject, limited tense variation of the verb, and a low degree of syntactic integration between the main clause and the complement clause (Diessel & Tomasello, 2001, pp. 106–107). They further asserted that *think*, contrary to *know*, was rarely negated in the children's speech. We also detected this asymmetric pattern in our data (Horbowicz et al., 2020). In a similar vein, emphasizing how an assumed sub-pattern (i.e., the negated pattern) of a construction can be critical to L2 development, Eskildsen (2018) found that *wh*-subclauses in the spoken L2 English of one learner were predominantly attached to the phrase *I don't know*, suggesting that it may be functioning as a stepping stone for syntactic development.

However, researchers have also found that some patterns appear to schematize rapidly and be productive from the start (Roehr-Brackin, 2014; Lesonen et al., 2020a). In a study of Finnish evaluative constructions with *tykätä* 'like' and *haluta* 'want', Lesonen et al. (2020a) found that development only to a very limited extent emerged from lexically specific patterns. Similarly, Roehr-Brackin (2014) documented that, among the German verbs *gehen* 'go' and *fahren* 'go by vehicle', only *fahren* followed a path from lexically specific to more productive abstract patterns. Roehr-Brackin argued that reliance on explicit knowledge related to cross-linguistic awareness, and instruction might have played a role. These findings imply that adult L2 development is probably not explainable by one learning mechanism alone (cf., for instance, the implicit/explicit debate; e.g., Ellis, 2015).

Summing up, even though the research suggests that the usage-based learning trajectory (i.e., from formula to full schematicity) is not always clear-cut in adult L2 learning, the evidence pointing out a significant role for semi-fixed utterances with one or more fixed and one open slot is convincing.

Although debated, formulaic language and fixed multiword expressions undoubtedly play a part in L2 learning and use, as also do semi-fixed patterns (Ellis, 2002, 2012). The debate has addressed the potential of formulas to trigger the expansion of patterns, for example, to what extent formulas must be analysed in order to contribute to further learning (cf. Wray, 2000; Eskildsen & Cadierno, 2007) and consequently the appropriateness of formulas as data relevant to the developmental process (e.g., Krashen & Scarcella, 1978). The functions and roles of formulaic language in L2 use and development are also ambiguous and multiple. Ellis (2012) argued that formulas may serve as lexical and phrasal ‘teddy bears’ that learners hold on to because of their communicative function, indicating that their functionality exceeds being the initial seeds of learning. As discussed above, the usage-based linguistics literature also makes it clear that linguistic representations vary in abstractness and that ‘old’ formulas and schematized productive patterns may be used side by side in an L2 (cf. Langacker, 2000).

2.2. SEMANTIC AND SYNTACTIC PROPERTIES OF EPISTEMIC VERBS IN NORWEGIAN

In Norwegian, epistemicity is mainly expressed lexically by modal verbs, epistemic adjectives, and adverbs, particles, and constructions with lexical verbs on which we focus in this study (Faarlund, Lie, & Vannebo, 1997; Horbowicz et al., 2020). The repertoire of epistemic verbs in Norwegian includes verbs claiming knowledge, like *vite* ‘know’, and verbs that decline to claim knowledge, like *tro* ‘think/believe’, *anta* ‘assume’, *gjette* ‘guess’, and *mistenke* ‘suspect’, among others (Goddard & Karlsson, 2008, p. 231). The latter verbs have also been labelled non-factive predicators, as they commit the speaker neither to the truth nor the falsity of the proposition, whereas verbs like *vite* are considered factive predicators (Lyons, 1977, p. 795). A recent new addition to the epistemic repertoire in Norwegian is *tenke* ‘think’. Traditionally this verb was reserved for reporting cogitation, but it can also be used to report a thought which is not epistemically grounded in either knowledge or feeling, for instance, a suddenly occurring thought (Goddard & Karlsson, 2008). During the past two decades, the frequency of *tenke* used in both epistemic and evaluative meaning has increased, and the collocation *tenker at* ‘think that’ has been growing in number (Language Council of Norway, n.d.). However, many consider the use of *tenke* in epistemic contexts to be incorrect, arguing that it is a direct copy of the English *I think*.

Apart from the verbs mentioned, the lexeme *synes* ‘think/be of the opinion’ requires some attention because of its similarity to *tro*. *Synes* has only an evaluative meaning and is used to express subjective opinions rather than

beliefs about objective facts, as is the case with *tro*. The semantics of *synes* is very close to *tro*, the only difference residing in the nature of the evidential component, which is knowledge in the case of *tro* and a feeling in the case of *synes* (Goddard & Karlsson, 2008).⁴ The existence of several basic-level verbs of thinking and the semantic distinction between *tro* and *synes* are specific to the Scandinavian languages (Goddard & Karlsson, 2008), and are as such often focused on in language classes and textbooks. The semantic boundaries between *tro*, *tenke*, and *synes* may be difficult to perceive. Consequently, a negotiation between different verbs that may fill the verb slot in epistemic constructions can be expected to be part of the development of L2 Norwegian (cf. Nordanger & Horbowicz, *in press*).

Epistemic verbs are transitive verbs, yet their syntactic properties depart from those of other transitive verbs in that the argument of an epistemic verb is restricted to either an anaphoric third person pronoun *det* ‘it’ or a finite clause (FC).⁵ In the latter case, the result is a complex clause consisting of a main clause with the epistemic verb as predicate and a subordinate clause that expresses the proposition modified by the main clause.⁶ There is, however, some variation in the FC complements that epistemic verbs may take: while verbs like *tro* can take only nominal FCs (with the explicit complementizer *at* ‘that’ or non-marked) as in example (1), verbs like *vite* can also be followed by *if*-clauses (in Norwegian the corresponding marker is *om*) or a variety of *wh*-clauses, introduced by a *wh*-complementizer (in Norwegian *hva*, *hvordan*, and others), as in example (2).

(1) Jeg tro-r at han komme-r i dag.
 I think-PRS that he come-PRS in day
 ‘I think that he is coming today’

(2) Jeg vet ikke om / hvorfor han komme-r i dag.
 I know-PRS not if / why he come-PRS in day
 ‘I don’t know if / why he is coming today’

The complement can also appear clause-initially when realized by the anaphoric pronoun *det* ‘it’ or a morphologically unmarked nominal sentence (i.e., without *at*). These instances result in inverted word order where the verb

[4] In Swedish, the corresponding lexemes are *tro* and *tycka*, but the semantic content is the same in Norwegian and Swedish (Goddard & Karlsson, 2008).

[5] In addition, some verbs, such as *vite* ‘know’, may appear without complement, performing distinct discursive functions (see also Section 4.2.2).

[6] Note, however, that Diessel and Tomasello (2001) have questioned whether all instances of epistemic verbs followed by an FC can be seen as complex biclausal utterances (cf. Section 2.1 on the initial formulaicity of verb constructions with *think* and *know*).

precedes the subject, as Norwegian is a place-holding verb-second language (cf. Karlsson, 2003):

- (3) Jeg tro-r det. / Det tro-r jeg.
 I think-PRS that / That think-PRS I
 'I think so'

When the object of the verb is a complement clause appearing clause-initially, the epistemic construction functions more like a tag structure, more loosely attached to the preceding clause than in the case of the regular word order with the subject preceding the verb and the argument (Diessel & Tomasello, 2001).

Another feature distinguishing epistemic verbs from other transitive verbs is that the former are highly collocable with egophoric subjects, that is, subjects referring to speech act participants (Dahl, 2000, p. 38). In Swedish, for instance, the co-occurrences of first person pronouns and verbs like *think* or *know* have been found to belong to the most frequent two-word combinations in spoken language (cf. Karlsson, 2006). Thus, one can also expect reduced variation in the subject slot in the construction in L2 use.

2.3. THE SCHEMATIZATION OF EPISTEMIC VERB CONSTRUCTIONS IN L2 NORWEGIAN

As discussed in Section 2.1, Lesonen et al. (2020a) have provided a framework for the investigation of the schematization process, reflecting how new linguistic material emerges and develops from an L2 perspective. By conceptualizing the schematization process as a continuum, they distinguished between four different steps moving from 'fixed lexically specific (formulaic)' to 'mostly formulaic' (with one open slot) and via 'semi-schematic' (with more than one open slot) to 'fully schematic, productive and variable'. This approach is adopted here, but in order to meet the specifics of the present study, we have made some adjustments to Lesonen et al.'s continuum (see Table 1). Constructions with one or more open slots are represented as one category of 'semi-schematic' and 'semi-fixed', and Lesonen et al.'s (2020a) category 'fully schematic' is relabelled 'highly schematic', since it is, in our use, verb-specific. Finally, we have included a space for 'schematic epistemic constructions' that represent productive non-verb-specific epistemic verb–argument constructions.⁷ Lesonen et al. (2020a) measured schematization by the degree of productivity operationalized as the variability of the specific main verb (including person and tense inflection) and the variability of complements in this construction. These criteria are also applied here for determining the level

[7] A fully schematic construction would be represented as SUBJ + VERB + OBJ.

TABLE 1. *Continuum of schematization for epistemic verb–argument constructions*

Type of construction	Lexically specific; fully fixed	Semi-schematic; semi-fixed; one or more open slot	Highly schematic, productive	Schematic epistemic construction
Examples from the data	<i>Jeg tror det er morsom</i> 'I think it is funny'	(i) <i>Jeg tror det er</i> + ADJ/NP 'I think it is' + ADJ/NP (ii) <i>Jeg tror</i> + FC 'I think' + FC	SUBJ + <i>tro</i> + NP/FC 'think' + NP/FC	SUBJ + EPISTEMIC VERB + argument

of schematicity of each specific epistemic verb–argument pattern in our data (i.e., constructions with, for instance, *tro* and *vite* as the main verb). We expect that the schematization process happens on two levels: by the gradual expansion of the construction frame to include more verbs, and by the gradual development of more productive and variable representations within each specific verb–argument construction, i.e., the degree of creative use manifested as slot variation and verb inflection. We assume that these developmental processes are running in parallel in adult L2 learners.

The study asks the following research questions:

RQ1: Which verbs occur in epistemic verb–argument constructions over time?

RQ2: How do epistemic constructions with *tro* and *vite* develop in the learners' L2 within the project period?

3. Data

The present study targets L2 development in a critical phase when the learners are expected to climb from the A2 level to B1–B2 (for a detailed description, see Council of Europe, 2001), i.e., from a communicatively functional to a more advanced and independent competence.

The study builds on spoken conversational data collected over the course of approximately four months (17 weeks). The participants were four adult learners, and the recorded conversations took place between each learner and their former teacher (Paulina), who is also one of the researchers. We recorded 8–9 conversations on various topics with each participant (see the Appendix for an outline of conversation topic, length, and number of finite verbs). The topics were consistent across participants in that the conversations of each project week (datapoint) revolved around the same topic, although there was a

TABLE 2. *The participants' biographical data. The participants' names have been anonymized.*

Participant	Number of conversations	Age	Time of residence	Reported first language	Partner	Habitual home language
Emilio	9	Early 20s	6 months	Spanish	–	–
Filipe	9	Late 20s	4 years	Portuguese	Norwegian speaking	English
Linda	8	Mid 30s	1 year	English	Norwegian speaking	English
Sofia	8	Early 20s	5 months	Hungarian	Norwegian speaking	English

new topic every week with the exception of two weeks when the participants were encouraged to describe a picture (see the Appendix). The data were transcribed in ELAN (Sloetjes & Wittenburg, 2008).

The participants were all enrolled in the same Norwegian class in a Norwegian institution for higher education. They had diverse linguistic backgrounds and varied in age and time of residence (see Table 2). All participants had some higher education and proficiency in languages besides their first language. Emilio and Linda were raised bilingually, but we choose not to specify their linguistic repertoire in full detail to protect their anonymity.

The data did not intentionally elicit epistemic constructions, although some conversation topics may have encouraged more use of epistemic markers than others, such as in project week 3 (see the Appendix) where the learners were told to reflect on what they thought was going on in a photograph. The token frequencies show a general increase in the use of epistemic markers, including verb–argument constructions with mental verbs, in this particular week. However, the general trend was that the use of epistemic markers increased throughout the project period (see Horbowicz et al., 2020).

4. Results

We extracted from the data all tokens of verb–argument constructions with verbs typically conveying epistemic meaning (see Section 2.2) and other verbs used with epistemic meanings, including those where the verb in standard Norwegian does not express epistemicity (mostly *synes*). Uses where the subject and/or the argument slot was not realized (typical examples being *vet ikke* ‘know not’ or *tror jeg* ‘I think’) were also included.⁸ The main selection

[8] The extraction was performed manually by both authors.

criteria were therefore (1) transitive epistemic verbs and (2) other transitive verbs conveying epistemic meanings. This approach allows us to capture and explore in detail the learners’ L2 repertoire of epistemic verb–argument constructions, including both target like and non-target like uses. The initial analyses detect the use of verbs in epistemic constructions across participants and over time. This part of the study responds to RQ1 and aims to uncover how the repertoire of verbs with epistemic meaning develops (Section 4.1). Subsequently, following Ellis and Ferreira-Junior (2009b), we narrow the scope and investigate the variation in each construction slot in constructions with the two most frequent verbs, *tro* and *vite*. This second part of the analysis addresses the obtained level of schematization to answer RQ2 and focuses on changes and development in the inventory of the whole *tro/vite* construction (Section 4.2). An overall ambition is to enable a closer look into the interaction between formulaicity and productivity in the learners’ epistemic verb–argument constructions.

4.1. THE REPERTOIRE OF EPISTEMIC VERBS

Table 3 demonstrates that the repertoire of verbs occurring in epistemic verb–argument constructions in general is limited in the data. Altogether, we find six verbs in use: *tro* ‘think/believe’, *vite* ‘know’, *synes* ‘think/be of the opinion’, *tenke* ‘think’, *ane* ‘sense’, and *mistenke* ‘suspect’, and there is no sign that the repertoire expands within the data collection period. On the contrary, all

TABLE 3. *The repertoire of verbs with epistemic meaning across participants and time*

Project week		1	3	5	7	9	10	12	16	17	SUM
Emilio	Tro	5	6	4	3	2	6	13	10	7	56
	Synes	1				1					2
	Tenke						1				1
	Vite	3	8	4		6	8	7	3	8	47
	Ane								1		1
Filipe	Tro	15	7	12	6	5	5	14	10	15	89
	Tenke				1						1
	Mistenke							1			1
	Vite	3	4	7	10	13	10	11	6	12	76
Linda	Tro	1	8	1	7	3	6	21		7	54
	Synes	1	4		1	1	3				10
	Vite	1	2	4	4	5	4	5		6	31
Sofia	Tro							1	1		2
	Synes		5		1	3	7	4	1	6	27
	Tenke		3		1				2		6
	Vite		8		4	9	10	10	7	10	58
	Ane		1								1

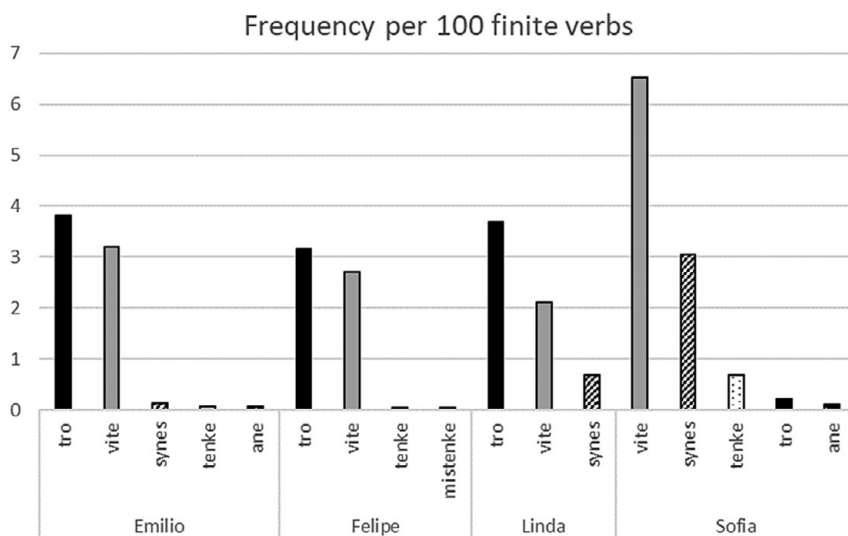


Fig. 1. The frequency of verbs with epistemic meaning per 100 finite verbs across participants.

participants employ no more than two verbs in the last recording. A calculation of the overall token number in the material reveals that two verbs are used much more frequently than the others, namely *vite* ($n = 212$) and *tro* ($n = 201$). Note, however, that *tro* occurs more frequently than *vite* in three of the participants' Norwegian (see Figure 1). *Synes*, which actually cannot have an epistemic interpretation in standard Norwegian (see Section 2.2), is ranked number three ($n = 39$). In line with previous research into children's and learners' verb use in specific constructions, the participants' use of epistemic verbs can be characterized as 'conservative' (cf. Ninio, 1999); that is, their repertoires are largely limited to a few highly frequent verbs (*tro*, *vite*, *synes*) which are also much more frequent than other occurring verbs (*tenke*, *ane*, *mistenke*). The dominance of *tro* and *vite* is not surprising, given the assuagingly high frequencies of both lexemes in spoken language (cf. Section 2.2; Dahl, 1997; Karlsson, 2006). It is worth noting, however, that *tenke*, which is used with increasing frequency with epistemic meaning in spoken standard Norwegian, is not very frequent in the learner data.⁹

Figure 1 reveals that Emilio and Felipe to a large extent show similar profiles, with two very frequent verbs and other verbs largely non-existent. Linda's and

[9] Unfortunately, no studies to date have addressed the frequency of *tenke* in epistemic expressions in spoken Norwegian, and it is beyond the scope of this paper to resolve this possible discrepancy between input frequency and L2 use.

Sofia's profiles, however, tell a different story. Like the other participants, Sofia's epistemic verb–argument constructions frequently contain *vite*, but unlike the others she seems to favour *synes* over *tro*. Sofia does not employ *tro* until late in the project period, and it occurs only twice. In contrast, she uses *synes* during the whole period with the semantic interpretation of *tro*. The frequency of her epistemic verb use resembles the Zipfian curve, with each verb being inversely proportional to its rank (cf. Ellis & Ferreira-Junior, 2009a), yet it seems that her epistemic constructions undergo the least development (see also Section 4.2). The calculation of produced finite verbs per conversation shows that Sofia produces the least language relative to the duration of the conversation (cf. the Appendix). Thus, she is a slow speaker who spends time assembling utterances, possibly implying that she is drawing on explicit resources, allowing her to avoid *tro* while choosing from a richer repertoire of lexical verbs.

Linda's use depicts some of the same characteristics as Sofia's, yet Linda's repertoire includes only three verbs. Even so, while Sofia uses *synes* as a direct replacement for *tro*, Linda's use is less consistent and more complex. A few of Linda's uses of *synes* occur alone with epistemic interpretation, but most are used in parallel with *tro*, and there seems to be an ongoing negotiation between *synes* and *tro* concerning the semantic and functional properties of the verbs, as in example (4).

- (4) Jeg tro-r jeg synes at hvis jeg kan jobbe i som psykologi kan jeg hjelpe dem
 I think- I think- that if I can work in as psychology can I help them
PRS PRS
 'I think I think that if I can work as a psychologist I can help them' (Linda, week 9)

The data suggest that Linda struggles with entangling the two semantic concepts, but according to our data, she no longer uses *synes* with epistemic meaning after week 12 (cf. Nordanger & Horbowicz, *in press*).

4.2. VERB–ARGUMENT CONSTRUCTIONS WITH *TRO* AND *VITE*

The following analyses focus on epistemic constructions with *tro* and *vite*, representing the two most frequent epistemic verb–argument constructions in the data. These sections respond to RQ2, asking how the learners' *tro/vite* constructions develop within this period of language learning, by mapping and tracing the linguistic material that may fill the different construction slots across learners and time.

4.2.1. *The subject slot*

Table 4 reveals that the level of variation found in the subject slot is generally low. The *tro* construction appears somewhat less varied than the *vite*

TABLE 4. *Subject slot occupants*

Subjects other than <i>jeg</i> 'I'	TRO		VITE	
	Frequency	Lexemes	Frequency	Lexemes
Emilio	2/56	<i>de fleste</i> 'most (people)', <i>man</i> 'one'	8/47	<i>vi</i> 'we' <i>du</i> 'you'
Filipe	0/89	–	8/76	<i>vi</i> 'we' <i>du</i> 'you' <i>folk</i> 'people' <i>alle</i> 'everybody'
Linda	3/54	<i>hun</i> 'she'	3/31	<i>du</i> 'you' <i>man</i> 'one' <i>mange</i> 'many (people)'
Sofia	0/2	–	1/58	<i>disse menneskene</i> 'these people'

construction, as all learners, except for Emilio, use more subjects with *vite* constructions than with *tro* constructions. The difference is perhaps most striking in Filipe's case. All of Filipe's *tro* constructions with a realized subject slot include the subject *jeg* 'I', while the subject slots in the *vite* construction show more lexical variation. Linda's profile is similar as she uses only *jeg* 'I' and *hun* 'she' with *tro*, but *du* 'you', *man* 'one', and *mange* 'many' with *vite*. Sofia employs *tro* only twice in the material, which makes it difficult to compare subject slots across these constructions.

In general, Table 4 shows that there is quite a strong consistency in the subject slot of the learners' epistemic constructions, and the *tro* construction subjects show less variation than those occurring with *vite*. These findings reflect Ellis and Ferreira-Junior's (2009b) results, discussed in Section 2.1, that the environment of the most frequent verbs also seems to show a high degree of predictability.

The data also include instantiations of the constructions where the subject slot is empty. There are in total 20 such occurrences in the *vite* construction, which is realized as *vet ikke* 'don't know': 7 in Filipe's data and 13 in Sofia's data. Those instantiations function as parenthetical remarks within a turn or as an answer to a question, and they are also fairly common in Norwegian L1 use. The *tro* constructions have three subject-less realisations, one in Emilio's data (*tror det* 'think that') and two in Filipe's data, both followed by a regular FC (*tror det er* 'think it is' and *tror du kan* 'think you can'). It is important to note that the subject-less *tro* constructions are found in weeks 1 and 3 only, while *vet ikke* appears throughout the data collection period in Sofia's and Filipe's data.

TABLE 5. *Arguments in tro constructions*

Argument-slot occupiers	FC complements with <i>at</i> 'that', % in parentheses	FC complements without <i>at</i> 'that', % in parentheses	NP-complement <i>det</i> 'it/that', % in parentheses	No argument, % in parentheses	Total number of <i>tro</i> constructions
Emilio	31 (55.3)	4 (7.1)	10 (17.8)	11 (19.8)	56
Filipe	0	42 (47.2)	9 (10.1)	38 (42.7)	89
Linda	11 (20.4)	31 (57.4)	4 (7.4)	8 (14.8)	54
Sofia	0	2 (100)	0	0	2

4.2.2. *The argument slot*

A broad share of the variation in the constructions is observed in the argument slot. As already noted, epistemic verbs can take a complement in the form of either a pronoun (NP) or a finite subordinate clause (FC). As the verbs *tro* and *vite* take different types of FCs as arguments (cf. Section 2.2), we will discuss these constructions separately.

Tro construction arguments. Table 5 shows that FCs make up approximately half of the complements in *tro* constructions, while NP complements represent between 7% and 18% (Sofia excluded).¹⁰ The subsequent discussion will thus first and foremost focus on the participants' use of FC complements.

There are striking differences among the participants in terms of the formal properties of the *tro* complements. The FC complements functioning as arguments to the *tro* constructions in Filipe's data are never marked by the complementizer *at*, and 19 out of 42 FCs start with *det er* 'it is', resulting in a highly frequent multiword expression *jeg tror det er* 'I think it is', which also seems to have a high degree of fixedness (see also Section 4.2.3). For instance, when the FC includes negation, it is often placed after the verb, as in example (5), which violates the target language norms demanding a preverbal position for sentence adverbials in a subclause.¹¹

- (5) jeg tro-r det er ikke så stor samling av folk
 I think-PRS it be-PRS not such big gathering of people
 'I don't think it is such a big gathering of people' (Filipe, week 17)

Example (6) illustrates another case of non-standard use of the *jeg tror* sequence, where the argument is realized by a minimal response particle:

[10] The remaining *tro* constructions are either *tror jeg*, used as a tag (attached at the end of a sentence or a phrase), or argument-less constructions (e.g., in self-correction instances).

[11] Note that examples such as (5) also occur in the native language use, although they do not follow the standard norms of Norwegian.

- (6) jeg tro-r nei
 I think-PRS no
 'I think no' (Filipe, week 3)

Such examples suggest that the *tro* construction is not fully schematized and occurs repeatedly as the *jeg tror X* pattern, where the X can be realized by either a different main clause or a phrase. Note also that the complementizer *at* actually occurs twice with other epistemic verbs in Filipe's data (*mistenke* 'suspect' and *tenke* 'think'), again strengthening the impression that the pattern with one open slot *jeg tror X* is highly fixed in his L2 Norwegian.

Contrary to Filipe, Emilio's FC complements in *tro* constructions are predominantly used with the complementizer *at*. Emilio also frequently employs *det er* as the initial sequence in the subclause. For example, *jeg tror at det er* 'I think that it is' occurs 7 times out of 29, and *jeg tror det er* 'I think it is' occurs twice, but there is in general considerable variation, both within the FCs following *tror* and within the other construction slots. This variation includes a variety of different subjects and verb forms, indicating that, unlike in Filipe's data, there is no dominating semi-fixed pattern (cf. Table 1). There also seems to be a stronger syntactic integration between the main clause and the FC, observable in the predominant use of the complementizer and in the temporal congruence of both main and subclause predicates, as in examples (7–8).

- (7) jeg aldri tro-dde at jeg ville føle meg på denne måt-en
 I never think-RET that I will-RET feel me on this way-DEF
 'I never thought that I would feel this way' (Emilio, week 7)

- (8) jeg trodde at det var sånn i alle land-e
 I think-RET that it be-RET such in all country-PL
 'I thought it was like that in all countries' (Emilio, week 12)

In general, Emilio's use of the *tro* construction appears more variable than Filipe's, which may indicate that the construction is more schematic (cf. Table 1; Lesonen et al., 2020a).

Whereas Filipe predominantly constructs subclauses without complementizers and Emilio predominantly uses *at*, Linda uses both. Among all four participants, she has the highest percentage of FC complements (see Table 4). Of Linda's 54 *tro* constructions, 8 are used without complement and 4 instances of *tro* occur with an NP complement (*jeg tror det* 'I think so'), while 42 constructions include an FC. Among the FCs, there are 20 occurrences of *jeg tror det er* 'I think it is', but both *SUBJ tror at det er* 'SUBJ think that it is' and *jeg tror at det er* 'I think that it is' occur. Overall, the sequence *det er* initiates more than 50% of the FCs in the *tro* construction. Linda's instantiations of the construction do not include tense variation in either the main clause or the FC, and the *det er* sequence is never interrupted by negation or adverbials. Even though there is variation in the *tro* subclauses, Linda also employs some

multiword expressions that appear highly fixed. The phrase *jeg tror det er veldig morsom* ‘I think it is very funny’ occurs 3 times in week 12, once the same week in the form *jeg tror det er morsom* ‘I think it is funny’, and twice in week 3 and 9 with *hun* ‘she’ filling the subject slot. Interestingly, the function of these phrases also seems consistent, although they are not easily comprehensible. They seem to indicate that something is funny, which would normally not be considered a question of epistemicity, but rather a state of opinion. Furthermore, *jeg tror det er veldig ADJ*, exemplified in example (9), occurs frequently in the second half of the data collection period.

- (9) jeg tror det er veldig viktig
 I think-PRS it be-PRS very important
 ‘I think it is very important’ (Linda, week 12)

Like Filipe, some of Linda’s *tro* constructions can be interpreted as semi-fixed patterns, possibly with formulaic bases, and perhaps also playing the role of ‘phrasal teddy bears’ (Ellis, 2012) in her spoken Norwegian.

Vite construction arguments. As Table 6 shows, a great proportion of the *vite* constructions in the data are argument-less. Among the participants, only Linda has more than 50% of uses with the argument slot occupied by an FC or an NP, while the corresponding number for Sofia is 18.9%. The majority of such argument-less constructions is represented by *jeg vet ikke* ‘I don’t know’, used as a parenthetical remark within the speaker’s turn or as a separate turn.

The *vite* construction allows for a greater variety of FC complements than the *tro* constructions (cf. Section 2.2). In line with the target language norm, the participants use both *at* ‘that’ and *wh*-pronouns (*hva* ‘what’, *hvor* ‘where’, etc.) as complementizers for *vet*, but they use *hvis* ‘if’ in contexts where the target norm would prefer *om* ‘if’.¹² The use of FC complements is consistent across all four participants of the study: the majority are *wh*-clauses predominantly following *jeg vet ikke*.¹³ Alternatively, *vite* constructions may take an NP argument: *jeg vet (ikke) det* ‘I know (not) that’ (see Table 6), but these are much less frequent in the present data than are FC clauses.

A clearer developmental dimension can be observed for *vite* constructions than for *tro* constructions. Throughout the project period, new

[12] In Norwegian, there is a distinction between the complementizer *hvis*, used in hypothetical sentences, and the complementizer *om*, used in embedded ‘yes/no’ questions, while English uses *if* in both contexts. That the participants in the study consequently use *hvis* for *om* may indicate that they rely on their English and/or mother tongue knowledge when constructing similar sentences in Norwegian.

[13] The share of *wh*-clauses following *jeg vet ikke* of all *wh*-clauses used in the *vite* construction is: 7 out of 10 for Emilio, 7 out of 12 for Filipe, 15 out of 16 for Linda, and 6 out of 6 for Sofia.

TABLE 6. *Arguments in vite constructions*

Argument-slot occupants	FC complements % in parentheses	NP <i>det</i> 'it/that' % in parentheses	(<i>Jeg</i>) <i>vet ikke</i> without argument, % in parentheses	Other instances without argument, % in parentheses	Total tokens of <i>vite</i> constructions
Emilio	19 (40.4)	3 (6.4)	21 (44.7)	4 (8.5)	47
Filipe	21 (27.6)	4 (5.3)	35 (46)	16 (21.1)	76
Linda	17 (54.8)	0	10 (32.3)	4 (12.9)	31
Sofia	10 (17.2)	1 (1.7)	45 (77.6)	2 (3.5)	58

complementizers appear in the FC complement used by the participants, as summarized below:

EMILIO: *at* 'that' (week 1) → *hvis* 'if', *hva* 'what' (week 3) → *hvordan* 'how' (week 10) → *hvilken* 'which' (week 12)

FILIFE: *hvordan* 'how', *hva* 'what', *hvis* 'if' (week 7) → *at* 'that' (week 9)

LINDA: *at* 'that' (week 1) → *hva* 'what', *hvor* 'where' (week 5) → *hvis* 'if', *hvorfor* 'why' (week 7) → *hvordan* 'how' (week 12) → *hvem* 'who' (week 12)

SOFIA: *hva* 'what' (week 3) → *hvis* 'if' (week 7) → *hva slags* 'what kind' (week 9)

Interestingly, all complementizers except *at* are used as a complement to *jeg vet ikke* 'I don't know', and a common denominator for all four participants is that *hva* 'what' is the first *wh*-complementizer used. Note also that all *wh*-complementizers follow a negated construction *jeg vet ikke*. This finding resonates well with Eskildsen's (2018) observation that *wh*-subclauses seem to be closely attached to *I don't know* in English. Of all participants, Linda shows the highest percentage of subclause arguments and the most gradual development of how different *wh*-complementizers are integrated into the *vite* construction (cf. Table 7).

Linda's first occurring *wh*-complementizers are the same as the first used *wh*-pronouns reported for Valerio by Eskildsen (2015). Note further that all of Linda's *wh*-complementizers are used repeatedly in the same context later in the period, with the exception of *hvordan* 'how', which appears only once in week 12. The L2 Norwegian data thus imply that *jeg vet ikke* may be functioning as a stepping-stone to developing syntactic complexity. In general, the *vet* construction in the material as a whole is heavily dominated by *jeg vet ikke* used either as an isolated multiword expression or as a point of departure for *wh*-subclauses.

TABLE 7. *Linda’s development and use of FC complements in the vite construction*

New complementizers per project week	1	3	5	7	9	10	12	17
Jeg vet FC	at (1)				–	–	–	(1)
Jeg vet ikke FC			hva (1)	(2)	–	(1)	(1)	–
Jeg vet ikke FC			hvor (1)	–	(1)	(1)	(1)	(1)
Jeg vet ikke FC				hvorfor (1)	(1)	–	–	(1)
Jeg vet ikke FC				hvis (1)	–	–	–	(1)
SUBJ vet-PRET ikke FC							hvordan (1)	–
							hvem (1)	–

4.2.3. *The learning trajectories of epistemic tro and vite constructions*

It is an overall goal of the present study to capture the process of schematization in L2 epistemic verb–argument constructions. Following Lesonen et al. (2020a), in Section 2.3 we proposed a continuum of four stages for the schematization of epistemic verb–argument constructions from fully fixed to abstract and schematic patterns. These stages are assumed to represent developmental steps (cf. Lesonen et al., 2020a). To this point, the results have focused on the linguistic material that may fill the different slots in the learners’ epistemic *tro/vite* constructions. However, to determine a construction’s level of schematicity, it is also important, as outlined in Section 2.3, to take into account the productivity of the construction, for instance, reflected in verb-slot variation. Another criterion is the extent to which basic construction patterns may be interrupted by adverbials and inversion as signs of a lesser degree of fixedness, or the degree of verb inflection. A highly productive, yet verb-specific, pattern could be represented as SUBJ + *vite/tro* + OBJ (cf. Table 1). The following paragraphs go through the participants’ *tro/vite* constructions.

For *tro* constructions, there are considerable individual differences among the three participants (Sofia’s use of *tro* is very limited, and she is therefore not included in the following discussion). Table 4 shows that subject-slot variation in *tro* constructions is low in general. A low degree of variation is, to varying extent, also characteristic of the whole construction for some of the learners. Whereas Filipe and Linda seem to rely on a highly fixed base for their *tro* constructions, Emilio has several uses where the verb is inflected for tense from week 5 onwards, implying that the construction is to some degree already represented as the abstract *tro* + OBJ (see example (10) and examples above).

- (10) jeg tro-dde at ikke sant det er
 I think-PRET that not right it be-PRS
 ‘I thought that, isn’t it right, it is’ (Emilio, week 5)

From approximately the middle of the data collection period, Emilio's *tro* constructions also include negation, as in example (7). In Emilio's data, negated *tro* occurs four times. This is not the case for Linda and Filipe, where negation appears only once. In Filipe's data, the argument slot following the negation is filled by an adverb rather than an FC, as in example (11).

- (11) jeg tro-r ik- ikke ennå men kanskje etterpå
 I think-PRS no- not yet but maybe afterwards
 'I think no- not yet but maybe afterwards' (Filipe, week 7)

Interestingly, Linda's one and only negated *tro* is followed by an FC with a referential NP as a subject rather than the usual *det* 'it', as illustrated in example (12).

- (12) jeg tro-r ikke norsk menneske-r ikke skjønne hvordan de kan kjøre
 I think- not Norwegian person-PL not understand how they can drive
 PRS
 'I think not Norwegian people don't understand how they can drive' (Linda, week 17)

Linda's data contain no instances with the inverted word order when the *tro* construction is used as a tag, while Filipe and Emilio make frequent use of this pattern throughout the data collection period. It seems therefore that Filipe's and Emilio's *tro* constructions are more productive than Linda's, yet for Filipe the variability is mainly visible in the word order, while Emilio also shows additional signs of productivity, such as the use of tense inflection and negation, discussed above and in Section 4.2.2.

All four learners show more variation in the subject slot for *vite* constructions, implying that this pattern is indeed used more productively than *tro* constructions (cf. Table 4). Otherwise, the participants once again show differences regarding the degree of variation observed in the construction. For instance, Sofia's *jeg vet ikke* sequences are never interrupted by other linguistic material. Instead, there are examples in the material where adverbs are attached sequentially to the *jeg vet ikke* pattern, as in example (13), while the target language normally requires adverbs to precede negation.

- (13) jeg vet ikke faktisk
 I know-PRS not actually
 'I don't know actually' (Sofia, week 3 and week 16)

Furthermore, compared to the other participants, Sofia's employment of FCs remains limited in both tokens and types throughout the data collection period. We interpret this as a possible sign of formulaicity and lack of development of the *jeg vet ikke* construction. Linda's *vite* constructions do not include adverbs, but her uses of *vite* constructions signal fixedness in other ways. For instance, Linda's *jeg vet ikke* pattern is never interrupted, and there are no occurrences in our material of tense inflection. However, as

shown in example (14), tense inflection occurs once when the subject is *mange* ‘many’, but never with *jeg* ‘I’.

- (14) mange visste ikke hvem de skal velge på
 many know-PRET not who they shall-PRS choose on
 ‘Many didn’t know who they should choose’ (Linda, week 12)

This is different in Filipe’s data. Although the *jeg vet ikke* pattern is very frequent, there are several examples of tense inflection and inverted word order. This syntactic flexibility and integration indicate that both Filipe and Linda may be on their way to developing an abstract pattern for *vite*, although Linda’s *jeg vet ikke* seems to be used as a fixed formula.

The learners’ use of epistemic *tro* and *vite* constructions do not easily lend themselves to uniform categorization. To sum up, however, on a scale from ‘fully fixed’ to ‘highly productive and fully schematic’ (cf. Table 1), Sofia’s epistemic *vite*-constructions appear closer to fixed expressions than Linda’s, which may be categorized as ‘semi-fixed’. Filipe and Emilio’s use of *vite* allows for more variation, hence resembling a ‘highly productive’ construction, and so do Emilio’s *tro* constructions. Filipe and Linda’s *tro* constructions, on the other hand, seem best described as ‘semi-fixed’. The nuances of these broad generalizations are discussed in detail below.

5. Discussion

In this study, we have followed four learners intensively through their second semester of Norwegian studies and documented the use of epistemic subject–verb–argument constructions throughout a period of four months. In response to RQ1, focusing on the employment and integration of lexical verbs with epistemic meaning, the data indicate that the repertoire of epistemic verbs is rather limited and does not increase during the data collection period. On the contrary, with time the participants use fewer verbs with epistemic meaning than they do at the beginning of the data collection period. This finding may at first glance seem to depart from Römer and Berger (2019), who found that the learners’ constructions became more varied as proficiency increased. However, we cannot assert if this may be a matter of differences in proficiency level, the developmental window investigated, the type of construction, or mode. For three of the participants, the repertoire of epistemic constructions consists of two highly frequent exemplars, which are predictably generic in meaning (Ninio, 1999; Ellis & Ferreira-Junior, 2009a; Römer & Berger, 2019). Still, due to the lack of new verbs being taken into use, we cannot assert whether those two verbs, *tro* and *vite*, lead the way for other verbs in our data. A data collection covering a longer period of time would probably have provided more answers. Moreover, it could have been interesting to pursue the idea of *tro*

functioning as a syntactic pathbreaker for other mental and perception verbs (cf. Diessel & Tomasello, 2001) with shared construction properties.

RQ2 addressed to what extent we could observe development in constructions with *tro* and *vite* within this period of L2 development with the overall goal of studying the process of schematization. We approached this question by a stepwise investigation of the different construction slots before addressing the developmental dimension for the construction as a whole. In line with Ellis and Ferreira-Junior's (2009b) results, our analyses demonstrate that there is a quite clear consistency in what may fill the subject and argument slots of *tro* and *vite*. The consistency is higher for the subject slot, and *tro* is complemented by a less varied argument slot than *vite*. The prototypical subject for both verbs is the first-person pronoun *jeg* 'I', yet *vite* does occur with other subjects in all participants' data, while *tro* shows less variation in this respect. Given the fact that mental verbs frequently appear with egophoric subjects in spoken language (see Section 2.2), we can conclude that the L2 use mirrors the frequency in the target language and that the lack of variation in the subject slot alone cannot be treated as a sign of lacking schematicity of the whole construction. At the same time, the argument slot is also less varied in the *tro* construction. The most common argument of the *tro* construction is the FC, yet in the data of Filipe and Linda, these FCs often do not seem syntactically and/or conceptually integrated with the main clause. As such, the sequence *jeg tror* followed by an FC can often be seen as a kind of epistemic marker, attracting attention or guiding the hearer's interpretation of the utterance rather than expressing a full-fledged proposition (Diessel & Tomasello, 2001, p. 97). This discursive functionality is somewhat reminiscent of what Ellis (2012) labelled 'teddy bears', or formulaic combinations of words performing distinctive communicative functions. Yet, more research is needed to ascertain the communicative role of such semi-formulaic epistemic patterns. As noted before, Eskildsen (2009) has suggested that the importance of learners' semi-formulaic construction patterns should not be underestimated, and that semi-schematic and semi-fixed patterns, also representing productivity and communicative functionality, may indeed be the endpoint of L2 learning. In light of recent research questioning if adult speakers draw on abstract schemas at all (Ambridge, 2020), this proposal is no less relevant. It is thus interesting to note that Filipe and Linda, the participants with the longest stays in Norway prior to the language course and the most fluent speech (cf. the Appendix), are the ones who rely on formulaic or semi-fixed patterns the most, thus showing similar linguistic behaviour, as noted by Eskildsen (2009, 2012, 2015, 2018). Lesonen et al. (2020a, 2020b) reported a similar finding: of four learners, the two learners with longer stays in Finland prior to language classes were less varied in their use of evaluative constructions than the learners who started the course immediately after arrival.

As indicated in Section 4.2.3, the present findings point out clear differences between the learners regarding the level of fixedness and schematicity. For all learners, the *vite* construction base prototypically includes negation, whereas the opposite is true for the *tro* construction. However, Sofia and Linda appear to be least flexible in terms of how they use the *jeg vet ikke* pattern. In their Norwegian, this pattern remains highly fixed throughout the data collection period. It is important to note that Linda's *jeg vet ikke* seems to have an increasingly productive argument slot, providing *jeg vet ikke* with the role of a stepping-stone for syntactic subclause development. During the data collection period, Linda attaches more and more different *wh*-clauses to *jeg vet ikke*. Similar developments are documented for Filipe and Emilio, who both include more *wh*-clauses within this period. For the *tro* construction, the picture is slightly different, as Filipe and Linda seem to form a group distinct from Emilio. Emilio's *tro* constructions demonstrate a high level of productivity and abstractness from early on (cf. Section 4.2), and his use of *tro* may be categorized as a 'highly productive pattern', while Linda and Filipe to a larger extent seem to rely on a fixed base *jeg tror* with an open slot corresponding to 'semi-fixed patterns' (cf. Table 1; Eskildsen, 2009, 2014, 2017; Lesonen et al., 2020a). This pattern remains frequent throughout the entire data collection period. In fact, negation occurs only once in Linda's data in the last project week (17), and neither Linda nor Filipe inflect *tro* for past tense. Conversely, Emilio negates *tro* to a larger extent, which we interpret as a sign of the construction's growing abstractness. Moreover, both Linda and Filipe use *jeg tror det er* 'I think it is' frequently, and Linda also repeatedly employs the multiword expression *jeg tror det er veldig (morso)* 'I think it is very (funny)'. Our data cannot confirm that these utterances are formulaic and fully fixed, but for both participants the usage frequency and lack of interrupting elements in these constructions imply that they may be classified as 'semi-fixed patterns'.

As stated in Section 2, our data cannot confirm that learning will indeed always start with formulaic patterns and develop into schematic ones, which is not the goal of the present study. In fact, except for Emilio, the participants' *tro/vite* constructions first and foremost seem to be used in a way that may best be described as semi-schematic within this period, and much of the development that occurs also falls within this category in that the development mostly concerns the inclusion of more complements. Linda's *jeg vet ikke* FC, reflecting a gradual integration of more advanced structures, is an illustrating example. Taking previous research into account (cf. Section 2.1), this finding comes as no surprise.

As noted in the brief interim summary in Section 4.2.3, although there are signs of emerging schematicity in terms of negation and tense inflection also in Linda's and Filipe's use of *tro* and *vite*, our data indicate quite clearly that

learners' constructions consist of formulaic and semi-fixed productive patterns that run in parallel. Additionally, constructions may also be used in a formulaic way and consist of both fixed and productive elements concurrently when the learners' level of proficiency increases. Finally, our data point out that there may be some substantial differences between newly arrived and experienced language class participants. Specifically, semi-fixed patterns seem to play a more central role in Linda's and Filipe's language than in Emilio's and partly Sofia's (in terms of verb repertoire). This paradox may be explained by the extent to which the learners draw on explicit resources. As noted by Lesonen et al. (2020a) and Roehr-Brackin (2014), some constructions seem to be productive and abstract from the start, possibly due to explicit knowledge. This may also explain Emilio's highly productive, flexible, and possibly schematic use of *tro/vite* constructions, as well as Sofia's verb repertoire, as mentioned above. This explanation is supported by the fact that they are both highly aware speakers who spend time assembling utterances, resulting in less produced language than Linda and Filipe relative to the time they speak (see the Appendix). These findings lead to the conclusion that, for adult learners, the usage-based learning trajectory also sometimes seems to be the other way around. In such cases, the developmental trajectory might not necessarily be a question of schematicity but of entrenchment, from highly flexible and productive analysed patterns to the parallel establishment and use of more fixed and formulaic routine language.

6. Conclusions

The ambition of the present study was to investigate the schematization process occurring in epistemic verb–argument constructions in post-initial phases of Norwegian L2 development. In line with previous research (cf. Eskildsen, 2014, 2017), we may conclude that all learners show signs of growing schematicity, but that these signs vary from individual to individual and from construction to construction, as also found in Eskildsen (2012, 2015, 2020) and Lesonen et al. (2020a). The different levels of schematicity associated with each verb-specific epistemic verb–argument construction point to the relevance of Tomasello's (2000) concept of 'verb island': there is no clear-cut uniform grammatical development across constructions; instead, each appears to develop following its own trajectory. There are also clear inter-individual differences regarding the reliance on formulaic patterns. The results suggest that the degree of drawing on explicit knowledge may be a crucial variable in whether the learners' L2 development will be dependent on semi-fixed and formulaic sequences. Concurrently, fixed multiword expressions with epistemic verbs appear to perform an important communicative function, allowing the speakers who rely on them to communicate more fluently and

possibly more pragmatically efficiently (cf. Diessel & Tomasello, 2001) in the L2. In line with previous research (Eskildsen, 2009, 2012, 2014, 2017, 2018; Lesonen et al., 2018), our analyses show that formulas may function as important stepping-stones in the development of more advanced syntax. A general conclusion to be drawn is thus that formulaic language deserves attention on all levels of L2 competence.

REFERENCES

- Ambridge, B. (2020). Against stored abstractions: a radical exemplar model of language acquisition. *First Language* 40(5/6), 509–559.
- Bartning, I. & Hammarberg, B. (2007). The functions of a high-frequency collocation in native and learner discourse: the case of French *c'est* and Swedish *det är*. *International Review of Applied Linguistics in Language Teaching* 45, 1–47.
- Bybee, J. (2006) From usage to grammar: the mind's response to repetition. *Language* 82, 711–733.
- Council of Europe (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Dahl, Ö. (2000). Egophoricity in discourse and syntax. *Functions of Language* 7(1), 37–77.
- Dabrowska, E. & Lieven, E. (2005). Towards a lexically specific grammar of children's question construction. *Cognitive Linguistics* 16(3), 437–474.
- Diessel, H. (2019). *The grammar network: how linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.
- Diessel, H. & Tomasello, M. (2001). The acquisition of finite complement clauses in English: a corpus-based analysis. *Cognitive Linguistics* 12(2), 97–141.
- Ellis, N. (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, N. (2006). Cognitive perspectives on SLA: the associative-cognitive CREED. *AILA Review* 19(1), 100–121.
- Ellis, N. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics* 32, 17–44.
- Ellis, N. (2015). Implicit AND explicit language learning: their dynamic interface and complexity. In P. Rebuschat (ed.), *Implicit and explicit learning of languages* (pp. 1–24). Amsterdam/Philadelphia: John Benjamins.
- Ellis, N. & Ferreira-Junior, F. (2009a). Construction learning as a function of frequency, frequency distribution, and function. *Modern Language Journal* 93(iii), 370–385.
- Ellis, N. & Ferreira-Junior, F. (2009b). Constructions and their acquisition: islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics* 7, 187–220.
- Eskildsen, S. (2009). Constructing another language: usage-based linguistics in second language acquisition. *Applied Linguistics* 30(3), 335–357.
- Eskildsen, S. (2012). L2 negation constructions at work. *Language Learning* 62(2), 335–372.
- Eskildsen, S. (2014). What's new? A usage-based classroom study of linguistic routines and creativity in L2 learning. *International Review of Applied Linguistics* 52(1), 1–30.
- Eskildsen, S. (2015). What counts as a developmental sequence? Exemplar-based L2 learning of English questions. *Language Learning* 65(1), 33–62.
- Eskildsen, S. (2017). The emergence of creativity in L2 English: a usage-based case-study. In N. Bell (ed.), *Multiple perspectives on language play* (pp. 281–316). Berlin: Mouton de Gruyter.
- Eskildsen, S. (2018). L2 constructions and interactional competence: subordination and coordination in English L2 learning. In A. Tyler & L. Huang (eds), *What is applied cognitive linguistics? Answers from current SLA research* (pp. 63–98). Berlin: Mouton de Gruyter.

- Esildsen, S. (2020). Creativity and routinization in L2 English: two usage-based case studies. In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer & R. Steinkrauss (eds), *Usage-based dynamics in second language development* (pp. 107–129). Bristol: Multilingual Matters.
- Esildsen, S. & Cadierno, T. (2007). Are recurring multiword expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics. In M. Nenonen & S. Niemi (eds), *Collocations and idioms 1. Papers from the first Nordic conference on syntactic freezes*. Joensuu, Finland, 19–20 May 2006 (pp. 86–99). Joensuu: Joensuu University Press.
- Faarlund, J. T., Lie, S. & Vannebo, K. I. (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Goddard, C. & Karlsson, S. (2008). Re-thinking THINK in contrastive perspective: Swedish vs. English. In C. Goddard (ed.), *Cross-linguistic semantics* (pp. 225–240). Amsterdam/Philadelphia: John Benjamins.
- Goldberg, A. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldberg, A. (2006). *Constructions at work*. Oxford: Oxford University Press.
- Goldberg, A., Casenhiser, D. & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics* 15(3), 289–316.
- Hammarberg, B. (2008) Konstruktioner som produkt och process: En studie av hur L1- och L2-talare utnyttjar «det är». *Nordand: Nordisk tidskrift for andrespråksforskning* 3(1), 79–107.
- Horbowicz, P., Nordanger, M. & Randen, G. T. (2020). Variabilitet og endring i utviklingen av epistemiske konstruksjoner i norsk som andrespråk: et dynamisk bruksbasert perspektiv. *NOA Norsk som andrespråk* 36(1), 5–38.
- Karlsson, S. (2003). Interactional uses for an epistemic marker: the case of 'Jag tycker'/'Tycker jag' in Swedish. *Melbourne Papers in Linguistics and Applied Linguistics* 3(1), 5–23.
- Karlsson, S. (2006). *Positioneringsfraser i interaktion. Syntax, prosodi och funktion*. Göteborg: Institutionen för svenska språket.
- Krashen, S. & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning* 28(2), 283–300.
- Langacker, R. (1987). *Foundations of cognitive grammar: theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (eds), *Usage-based models of language* (pp. 1–64). Stanford, CA: CSLI Publications.
- Language Council of Norway (n.d.). Å tenke i tide og utide. Online <<https://www.sprakradet.no/svardatabase/sporsmal-og-svar/a-tenke-i-tide-og-utide/>>.
- Lesonen, S., Steinkrauss, R., Suni, M. & Verspoor, M. (2020a). Lexically specific vs. productive constructions in L2 Finnish. *Language and Cognition* 12(3), 526–563.
- Lesonen, S., Steinkrauss, R., Suni, M. & Verspoor, M. (2020b). Dynamic usage-based principles in the development of L2 Finnish evaluative constructions. *Applied Linguistics* doi: 10.1093/applin/amaa030.
- Lesonen, S., Suni, M., Steinkrauss, R. & Verspoor, M. (2018). From conceptualization to construction in Finnish as an L2. *Pragmatics and Cognition* 24(2), 212–262.
- Lieven, E., Salamo, D. & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: a usage-based analysis. *Cognitive Linguistics* 20(3), 481–508.
- Lieven, E. & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. Ellis (eds), *Handbook of cognitive linguistics and second language acquisition* (pp. 168–198), New York/London: Routledge.
- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language* 26, 619–653.
- Nistov, I, Gustafsson, H. & Cadierno, T. (2018). Bruksbaserte tilnærminger til andrespråklæring. In A-K. H. Gujord & G. T. Randen (eds), *Norsk som andrespråk: Perspektiver på læring og undervisning* (pp. 107–132). Oslo: Cappelen Damm Akademisk.

- Nordanger, M. & Horbowicz, P. (in press). Ustabilit og uhandgripelig? En metodologisk diskusjon om dynamisk systemteori som tilnærming til variabilitet i andrespråksutvikling. *NOA Norsk som andrespråk*.
- Ochs, E. (1996) Linguistic resources for socializing humanity. In J. J. Gumperz & S. Levinson (eds), *Rethinking linguistic relativity* (pp. 407–437). New York: Cambridge University Press.
- Roehr-Brackin, K. (2014). Explicit knowledge and processes from a usage-based perspective: the developmental trajectory of an instructed L2 learner. *Language Learning* 64(4), 771–808.
- Römer, U. & Berger, C. M. (2019). Observing the emergence of constructional knowledge: verb patterns in German and Spanish learners of English at different proficiency levels. *Studies in Second Language Acquisition* 41(5), 1089–1110.
- Römer, U., Skalicky, S. C. & Ellis, N. (2020). Verb–argument constructions in advanced L2 English learner production: insights from corpora and verbal fluency tasks. *Corpus Linguistics and Linguistic Theory* 16(2), 303–331.
- Sloetjes, H. & Wittenburg, P. (2008). Annotation by category-ELAN and ISO DCR. *6th international Conference on Language Resources and Evaluation (LREC 2008)*. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. Online <<https://archive.mpi.nl/tla/elan>>.
- Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11(1/2), 61–82.
- Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
- Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics* 21(4), 463–489.

Appendix

Overview of the data, including conversation topics and recording lengths, measured in time and number of finite verb contexts (FV contexts).

Project week	Topic of the conversation	Emilio		Filipe		Linda		Sofia	
		Time	FV contexts	Time	FV contexts	Time	FV contexts	Time	FV contexts
1	Christmas holidays	16:04	143	12:44	219	10:48	156	11:52	53
3	Picture description: King Olav goes skiing	20:09	123	10:43	96	11:57	156	11:33	67
5	A place you like	20:40	147	24:09	393	08:45	150	-	-
7	Winter holidays	16:52	99	22:44	353	12:39	179	11:21	62
9	Choosing education paths	16:59	133	24:01	284	15:03	183	15:20	100
10	Picture description: King Olav goes skiing; talking to strangers	28:56	208	20:12	224	12:18	172	19:45	156
12	Do we need the Easter holidays?	27:43	168	26:33	373	19:57	273	18:04	103
16	Talk about the essay you wrote for the course	29:57	242	32:26	448	-	-	23:25	151
17	Recapitulation of the course and plans for the future	27:19	204	30:58	420	14:49	196	27:12	196
SUM		204:39	1467	204:30	2810	106:16	1465	136:52	888