# THE EARLY EXPRESSION OF (UN)CERTAINTY IN TYPOLOGICALLY DIFFERENT LANGUAGES: EVIDENCE FROM RUSSIAN, ESTONIAN AND HEBREW

Victoria V. Kazakovskaya[1],
Reili Argus, Sigal Uziel-Karl

**Abstract.** The article discusses the early means of expressing the epistemic notions of certainty and uncertainty (epistemic markers) in three typologically-different languages – Russian, Estonian and Hebrew. The results are based on an analysis of naturalistic speech samples of 9 typically-developing monolingual children, recorded from ages 1;3 to 6;2. The emergence of epistemic markers in child speech is compared to child-directed speech. Acquisition of means of expressing epistemic modality starts to develop at the end of the second year of life with the marking of uncertainty. The findings indicate an expansion of epistemic evaluation from objective situations in the physical world to the mental world. A comparison of child speech with the input reveals that both the frequency of a marker in the target system and the degree of epistemic semantics influence its emergence and development. Differences between languages mostly concern the frequency of usage of epistemic markers and the degree of epistemic evaluation children start with.

**Keywords**: Epistemic modality, first language acquisition, child-directed speech, Russian, Estonian, Hebrew

## 1. Introduction

It is commonly acknowledged that language is not just phonetics, lexicon and grammar, but is also a culturally-organized system. To become a native-like speaker of a language, a person has to acquire the cultural aspects of the language, e.g. when it is appropriate to use certain words or expressions, how to address others or make requests, etc., in addition to learning pronunciation, vocabulary and grammar. In different cultures, speakers use language differently to reflect social and cultural norms. For example, different languages have different means of expressing politeness.

In analyzing the relationship between language and cultural conceptualizations, the notions 'cultural schema' and 'cultural category' have been used as key terms (Sharifian 2017: 7). Cultural schemas and subschemas capture beliefs, norms, rules, and expectations of behavior as well as values relating to various aspects and components of experience. Cultural schemas are instantiated in many aspects of language. They capture encyclopaedic meaning that is culturally constructed. Cultural schemas may also provide a basis for pragmatic meanings: the knowledge underpinning the enactment and uptake of speech acts. This kind of knowledge has been assumed to be culturally constructed and therefore shared, that is, largely captured in such schemas (ibid: 33). In turn, cultural categories and subcategories are those culturally-constructed conceptual categories (colors, emotions, attributes, foodstuffs, kinship terms, events, etc.) that are primarily reflected in the lexicon (ibid: 26). Specifically, Nishida (1999) distinguishes eight primary types of schemas for social interaction (e.g. fact-and-concept, strategy, role schemas). Two of them may be relevant for the expression of epistemic modality, which is the focus of this chapter: i.e. context schemas, including knowledge about situations and associated appropriate behavior, and emotion schemas, including information about affect and evaluation. The latter are activated via their association with other schemas.

Epistemic modality is one of the domains of modality along with its deontic and dynamic domains. This category demonstrates the speaker's attitude and indicates his 'degree of confidence in a proposition' (Boye 2016: 117), i.e. his/her degree of certainty or uncertainty. Epistemic modality refers to the world knowledge of an individual based on opinion (Nuyts 2016: 38)[2]. However, our world knowledge and the way in which we refer to the factual status of a proposition are shaped by society. That is why the expression of epistemic semantics goes beyond the communicative function of language to have cultural significance. The appropriate behavior (context schema) may refer to how people in different cultures can speak about their knowledge of a situation and their attitudes towards it, including epistemic evaluation. Traditions of speaking about the degree of confidence of a proposition may vary cross-culturally. Moreover, the way we express our subjective opinion varies depending on whether we are speaking to adults (adult-directed speech, *ADS*) or children (child-directed speech (*CDS*, *input*)) and changes as children grow up. Investigations devoted to input in different languages and cultures provide some evidence for its variability (e.g. Ochs, Schieffelin (eds.) 1986, Lieven 1994) including the social economic status of families (Hoff 2006). However, the acquisition of epistemic modality has so far only been studied fragmentarily (Choi 2006, Matsui 2014, Kazakovskaya, Argus 2016) and most research has concentrated on deontic and dynamic domains (e.g. Hickmann, Bassano 2016, Stephany, Akcu-Koç (eds.) *forthcoming*).

---

[2]   A category very close to epistemic modality is evidentiality, expressing the source of information. Previous approaches that did not have access to the wealth of evidentials in the world's languages regarded these categories as overlapping or subsumed under a wider category of epistemic modality (Palmer 1986, Kiefer 2000). A logical connection between them exists, such that evidentiality asserts evidence, while epistemic modality evaluates evidence (de Haan 2005), but the categories can still be separated. In particular, in Estonian the use of an evidential *-vat* in marking the source of information is optional and dependent on various pragmatic conditions, whereas in Russian evidentiality is not grammaticalized.

The present paper focuses on the development of early (initial) expressions of epistemic semantics of certainty and uncertainty (hereinafter *(un)certainty*), i.e. epistemic markers (*EM*), in child speech (*CS*) as compared to CDS in 3 typologically-different languages – two contact ones, Russian (East Slavic, inflecting) and Estonian (Finno-Ugric, mostly agglutinating)[3], and Hebrew (Semitic, inflecting). These particular languages were selected since they are sufficiently different but at the same time share certain linguistic features, enabling the present authors to study and evaluate the extent to which typological and cultural factors might affect the early emergence of EM, as compared with structural factors like morphology and syntax.

The paper aims to explain the acquisition trajectory of EM with reference to the Russian, Estonian, and Israeli cultures as well as to the structural characteristics of the languages. Although all three countries are multinational, one can metaphorically define Russia and Estonia as mainly 'monocultural' countries, whereas Israel is a cultural 'melting pot'. However, although Russia might be described in such a way, it does geographically join Europe and Asia and their numerous nationalities, language dialects, and, consequently, various linguocultural traditions (which makes the description of a 'collective' linguocultural portrait of Russia difficult). In turn, Estonia has greater contact with other languages and cultures than other Nordic countries, e.g. Finland. Compared with Finnish-speaking people, Estonians are characterized by rapid speech with a frequent switching of turns, and relatively frequent double speaking (Pajupuu 1995). As for Hebrew, as far as conversational style and language are concerned, an outsider will find Israelis very direct and blunt, as Israelis are quite opinionated, at times argumentative, and tend to show a high level of confidence, which may have to do with the fact that almost all Israelis have military experience.

---

[3]   The North-West region of Russia is their zone of principal contacts.

With the purpose of showing the main linguistic and socio-cultural factors affecting the acquisition of epistemic modality (viz. certainty and uncertainty) in the languages under observation, we describe the means for expression of this category (§2), mention in what follows we its previous investigations in CS (§3), present the corpora analysed and method (§4), describe the main tendencies in acquisition of EM (i.e. developmental findings) (§5) and account for them in relation to linguistic and socio-cultural factors (§6).

## 2. Epistemic modality in Russian, Estonian and Hebrew

The present section offers an overview of the means for expressing the semantics of (un)certainty in the languages under investigation, focusing on epistemic subdomains, their prototypical EM, functions, and stylistic peculiarities.

### 2.1. EXPRESSION OF CERTAINTY/UNCERTAINTY

In Russian (*RUS*), epistemic meanings of (un)certainty are mainly expressed via parenthetical modal words (*navernoe* 'probably'[4], *konečno* 'of course', *možet byt'* 'maybe') and modal particles (*vrode by* '(it) seems, it looks as if'), syntactic constructions, viz. clauses with mental verbs like *dumat'* 'think', *sčitat'* 'consider', *znat'* 'know' (*Ja znajau, (čto) p*[5] 'I know (that) p') and special intonation (Vinogradov 1947, Russkaja Grammatika 2005, Birjulin, Kordi 1990, Beljaeva 1990 among others). Modal words and their collocations are considered as the basic (prototypical) means of expressing this category. Parenthetical modal words can be freely placed in the sentence: *konečno* 'of course' p (initial position) or p, *možet byt'* 'maybe' (final position). Moreover, in colloquial dialogues, they are often used in responses within adjacent pairs, particularly, in question–reply

---

[4]  Here and in the article as a whole, only some illustrative examples will be given.
[5]  By *p* we mean a proposition.

units: *Vaš syn Sandro podderživaet vaši dizajnerskie sposobnosti?* 'Does your son Sandro support your designing abilities?' – **Konečno** 'Of course' (www.ruscorpora). For RUS colloquial speech, reduced forms of parenthetical modal words (*naverno* from *navernoe* 'probably', *možet* from *možet byt'* 'maybe') as well as most modal particles are typical.

Prototypical means for expressing epistemic modality in Estonian (*EST*) are also lexico-grammatical. They include parenthetical modal words and sentences (**Selge (on), p** '(It is) clear, that p'), non-parenthetical modal adverbials and particles (*kindlasti* 'certainly'), predicates and predicatives (**Ma usun, p** 'I believe that p', *See* **on kindel,** *p* 'It's certain that p'). Non-prototypical (syntactic) means include constructions with a modal verb in 3ʳᵈ person singular (**Paistab,** *p* 'It seems that p') and a mental verb in *da*-infinitive (**Võib arvata,** *p* 'It can be believed that p'), subordinate clauses (*Ta,* **nagu teame,** *ei tule* 'She, as we know, will not come') or infinite constructions (*Ta* **minu teada** *ei tule* 'She, as I know, will not come') (Erelt 2017). Modal particles form an increasing and changing set of words (Valdmets, Habicht 2013). All adverbs can be used in colloquial speech and only *äkki* 'maybe' is not used in more formal contexts. *Võib-olla* can be used in formal contexts, whereas *äkki* is more suitable for informal ones. There are also results showing that *äkki* has a lower degree of uncertainty than *võib-olla*. However, only younger people evaluated it such; older people think that it is in the middle of the scale of uncertainty (Tüüts, Argus 2016).

In Hebrew (*HEB*) expressions of (un)certainty include modal adverbs, mental and modal verbs, and modal adjectives. Epistemic adverbs form the largest category of expressions of (un)certainty (*betax* 'sure', *bevaday* 'certainly', *itaxen* 'possibly', *ulay* 'maybe'). Most adverbs appear at the beginning of a sentence, followed by a complementizer and a subordinate clause (*Yitaxen she yered geshem hayom* 'It is possible that it'll rain today'). Verbs used to express epistemic notions include mental verbs, the most common of which are

*yode'a* 'know' and *xoshev* 'think' (*Ani **xoshev** she carix lalexet lishon* 'I think that it is necessary to go to sleep'), and the modal verb *yaxol* 'can' followed by the copula 'be' in the infinitive *yaxol ihyot* 'maybe, perhaps' used to express uncertainty. Adjectival modals include the epistemic adjectives *asuy* 'is likely to' and *alul* 'is liable to' followed by the copula 'be' in the infinitive used to express positive or negative probability, respectively (*Ze asuy lehacliax* 'This is likely to succeed', *Ze alul lehishaver* 'This is liable to break') (Berman 2011, 2014, Coffin, Bolozky 2005, Dromi 1980)[6].

## 2.2. EPISTEMIC SUBDOMAINS

The semantic structure of epistemic modality is represented by two domains (viz. certainty and uncertainty) which, in turn, might be divided into subdomains, according to the degree of the speaker's confidence, running from high to low. A degree analysis was conducted based on the ADS data. The general list of prototypical EM (see Table 1 below) – which are mostly (parenthetical) modal words – along with the degree of their semantics, is based on grammar and other representative studies (e.g. Panfilov 1977, Beljaeva 1990 for RUS, Kehayov 2008, Erelt 2017 for EST).

## 2.3. PRAGMATIC (DISCOURSE) FUNCTIONS OF EM

In all three languages EM may fulfil primary and secondary functions. The primary functions are related to the implementation of the epistemic value, i.e. to the qualification of information in terms of reliability.

---

[6]   In HEB epistemic and deontic modality do not always show a clear structural distinctiction; thus, 'certainty' and 'necessity' share many modal expressions: *carix* 'ought to', *xayav* 'obliged to', *muxrax* 'must'.

**Table 1.** Epistemic subdomains and their prototypical EM[7]

| Semantic domain | Degree of expressing semantics | RUS | EST | HEB |
|---|---|---|---|---|
| Certainty | High ('categorical certainty') | *konečno* 'of course' *dejstvitel'no* 'really' *na samom dele* 'actually' *pravda* 'truly' lit. 'truth' | *kindlasti* 'certainly' *tõesti* 'really' *muidugi* 'of course' *loomulikult* 'naturally' *kahtlemata* 'undoubtedly' | *betax* 'sure' *be'emet* 'really' *kamuvan* 'of course' |
| | Non-high (middle and low; 'problematic certainty') | *kak budto* 'apparently, it would seem' *vrode by* '(it) seems, it looks as if' (coll.) *vrjad li* 'hardly, unlikely (coll.)' *kažetsja* '(it) seems' | *paistab, et* '(it) seems that' *tegelikult* 'actually' | *xayav lihyot* 'must be' *carix lihyot* 'should be' |
| Uncertainty | Low | *po-moemu* 'in my opinion' *kažetsja* '(it) seems'[8] *požaluj* 'very likely, perhaps' | *ilmselt* 'apparently' *arvatavasti* 'probably, supposedly' | *kanir'e* 'apparently' *amur* 'supposed to' |
| | Middle | *možet byt'* 'maybe' *možet* 'may' *požaluj* 'very likely, perhaps' | *võib-olla* 'maybe' *vist* 'probably' *äkki* 'maybe (coll.)' *võimalik* '(it is) possible' | *ulay* 'maybe' *yitaxen* 'perhaps' *yaxol lihyot* 'could be' |
| | High | *navernoe* 'probably' *vidimo* 'apparently' *očevidno* 'obviously, clearly' | *vaevalt* 'hardly' | *loh yitaxen* 'perhaps not'[9] *loh yaxol lihyot* 'impossible' |

---

[7]   It is not possible to translate all EM in a very exact way, nor to present an exhaustive list.

[8]   Utterances without any EM express so-called zero (implicit) certainty; they are not considered in this paper.

[9]   Due to the semantic graduality of epistemic modality, the adjacent divisions of this scale partially overlap: e.g. *kažetsja* 'it seems' is interpreted as a low degree of uncertainty and, at the same time, as a problematic certainty (depending on the context) (Kazakovskaya, *forthcoming*).

In RUS the secondary functions relate to the actualization of the speaker's figure (e.g. Bulygina, Šmelev 1992, Jakovleva 1994, Kazakovskaya 2011). In this case epistemic value is neutralized completely or partially and additional semantic and/or pragmatic nuances appear in the discourse. These are, for instance, underlining, or even exaggeration of the primary epistemic semantics within the so-called modal agreement, a clarification of the relationship of motivation in the situation of inference, construction of a hypothetis, a polite way of expressing the speaker's opinion which is characteristic of intelligent, educated people (e.g. 'author's *we*'), and expression of irony.

In EST, the secondary functions of EM are to express evidentiality (indirect information is considered to be evaluated epistemically and to be less certain than direct (Müürsepp 2015)) and politeness by using uncertainty markers to express the speaker's opinion. The epistemic value of different EM depends heavily on the cultural context: particularly, the uncertainty marker *arvatavasti* 'probably, supposedly' can be used to express different levels of this semantics in a neutral and ironic context (Tüüts, Argus 2016).

In HEB, EM (specifically, certainty) in secondary functions have a mitigating effect, serving to reduce the level of certainty of a proposition in which they appear, or to express irony or disbelief.

Thus, in all three languages, the prototypical means of epistemic modality belong to the lexico-grammatical sphere. However, in RUS their repertoire is wider and more diverse than in EST and HEB. EM have free placement within sentences, some stylistic peculiarities and fulfil different pragmatic functions. In RUS, the number of EM expressing certainty and uncertainty is almost equal, while in EST, uncertainty markers prevail and EM of middle degree are much more numerous (Kazakovskaya, Argus 2016). One can conclude then, that the so-called cultural schema of expressing (un)certainty in the three languages under observation looks quite similar, where the primary functions of EM are concerned, and bigger differences occur in secondary functions, in terms of diversity and frequency.

## 3. Previous studies on the acquisition of epistemic modality

An investigation into the acquisition of epistemic modality can be informative with respect to the development of attitudes in children and their socialization. To date, the acquisition of this category in typologically-different languages (including the ones under observation) has only been studied fragmentarily (see Choi 2006, Matsui 2014, Kazakovskaya, Argus 2016). Interestingly, in the languages investigated so far, children start to express their epistemic evaluation with marked certainty (Bassano 1996, Hickmann, Bassano 2013, Noveck et al. 1996). However, comparative pilot studies, based on RUS and EST, did not confirm this priority (Kazakovskaya, Argus 2016).

A few references to this process in RUS can be found in studies (Gvozdev 1949, Stoljarova 1992, Kazakovskaya 2011, 2016, *forthcoming*) which are mostly based on diary data and do not contain any examination of CDS. These observations have shown that epistemic modality emerges later and is less frequently expressed than deontic modality. Markers of deontic modality emerge before 2;0 (Gvozdev 1949, Wiemer 1992, Oficerova 2005), whereas EM usually occur after this age. It was observed that both the emergence of EM and their repertoire show variability across children and the list of certainty markers is smaller than that of uncertainty (Kazakovskaya 2011). A few experimental studies have been conducted with older children (aged 6 to 15 years) (Ovčinnikova et al. 1999, Krauze 2004, Sedov 2004).

In EST, deontic modality is also acquired first; children first start to use requests at approximately 1;5–1;7 (Argus, *forthcoming*), while first EM emerge around the age of 2;0 (Kazakovskaya, Argus 2016). EST children acquire markers of uncertainty first and they are focused on the existence or non-existence of an object referred to by a noun or for referring mostly to the possibility of some kind of general situation.

Research on the acquisition of modality in HEB has focused mainly on dynamic and deontic modality (for a review see Uziel-Karl, *forthcoming*). Epistemic expressions are acquired one to five months later than deontic ones (ibid.). As for school age children, Berman (2014) notes that developmentally, 9 to 12-year-old children rely largely on deontic modality in their writing, referring to socially determined prohibitions or prescriptions. In contrast, from adolescence onwards, writers shift to more individual epistemic attitudes expressing possible or probable future contingencies. This shift is taken to be a function of increased age and literacy as evident in many languages, with HEB speaker-writers making use of the language-particular repertoire of formal devices available to them for this purpose.

To conclude, in the languages under observation, domains of deontic and dynamic modality are acquired earlier than epistemic ones. In RUS and EST, children start the marking of epistemic modality with uncertainty.

## 4. Data and method

### 4.1. PARTICIPANTS AND DATABASE

We analyzed longitudinal naturalistic speech samples of 9 typically-developing monolingual children (Table 2). They were video recorded in 'caregiver–child' spontaneous interactions in different settings (meal or play time, storytelling, etc.) to prevent bias. The recordings were transcribed and coded using CHILDES (MacWhinney 2000). Although the children's ages at the beginning of the observation period differ, all the children use utterances with MLU ≥ 1.

**Table 2.** Participants and database

| Language | Child (gender) | Length of recordings (hours) | Age | Child's MLU |
|---|---|---|---|---|
| RUS[10] | Liza (F) | 38 | 1;6–3;0 3;5–4;1 | 1.1–4.5 |
| | Filipp (M) | 32 | 1;5–2;8 | 1.8–4.7 |
| | Kirill (M) | 9 | 1;7–2;6 2;9–3;0 4;2–6;2[11] | 2.0–9.9 |
| EST | Andreas (M) | 69 | 1;7–3;1 | 1.1–5.0 |
| | Martina (F) | 16.5 | 1;3–3;1 | 2.3–5.7 |
| | Linda (F) | 36 | 1;3–4;6 | 1.6–3.1 |
| HEB | Lior (F) | 32 | 1;5–2;9 | 1.0–3.5 |
| | Smadar (F) | 15 | 1;6–2;4 | 1.5–4.5 |
| | Leor (M) | 33 | 1;9–3;0 | 2.0–3.7 |

EM were isolated and coded for epistemic domain and degree of confidence (see Table 1), part-of-speech, and type of utterances (non-interrogative/ interrogative).

### 4.2. DATA ANALYSIS

The data were analyzed quantitatively and qualitatively for each language. We examined the following: the age of emergence and initial features of EM; the development of each epistemic domain, focusing on growth and diversity of the EM repertoire, their degree of confidence and pragmatic functions; relations between CS and CDS for the aspects mentioned above and the similarities and differences in epistemic development between the children.

We can connect our expectation concerning the cultural aspect of this phenomenon in L1 with the frequency and diversity of EM in 'caregiver–child' communication, their pragmatic functions and contexts of usage.

---

[10]   For this investigation, the North-West region, which is close to Estonia.
[11]   During this period of observations 8 months were recorded.

## 5. Results

### 5.1 GENERAL RESULTS

This section describes findings which relate to the number of EM in each language relative to the number of utterances per participant. As Table 3 below shows, in all three languages the number of EM (and accordingly epistemically marked utterances) is quite low.

**Table 3.** Epistemically-marked utterances in CS and CDS

| Language | Corpus | CS | | CDS | |
|---|---|---|---|---|---|
| | | Total utterances | Utterances with EM (%[12]) | Total utterances | Utterances with EM (%) |
| RUS | Liza | 9277 | 34 (0.4) | 17502 | 291 (1.7) |
| | Filipp | 8546 | 13 (0.2) | 13009 | 61 (0.5) |
| | Kirill | 6261 | 112 (1.8) | 8472 | 337 (3.9) |
| EST | Andreas | 31929 | 51 (0.2) | 21271 | 380 (1.8) |
| | Martina | 3378 | 14 (0.4) | 3822 | 157 (4.1) |
| | Linda | 8545 | 102 (1.2) | 10864 | 354 (3.2) |
| HEB | Lior | 6691 | 33 (0.005) | 7778 | 147 (0.02) |
| | Smadar | 3753 | 18 (0.005) | 3363 | 29 (0.009) |
| | Leor | 7234 | 16 (0.002) | 7935 | 52 (0.007) |

Figures 1–3 show the distribution (in percentages) of EM denoting (un)certainty in each language. All corpora show a higher percentage of uncertainty markers, which are more dominant in CS than in CDS. This tendency seems to be most prominent in the speech of RUS Liza, EST Martina, and HEB Lior.

---

[12]  Frequency of epistemically-marked utterances relative to the total number of utterances for each dyad.
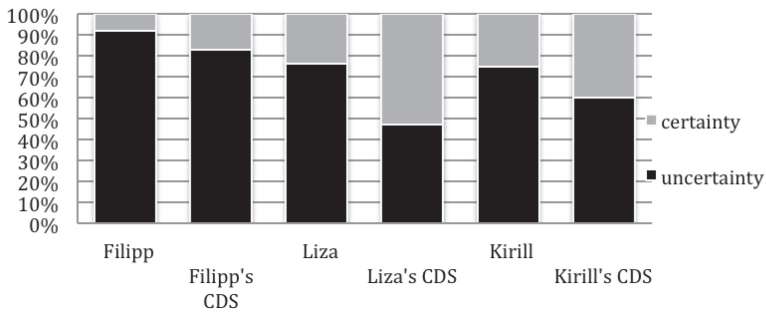
**RUS: Certainty *vs*. Uncertainty**



**Figure 1.** Distribution of EM (tokens) in RUS

**EST: Certainty *vs*. Uncertainty**
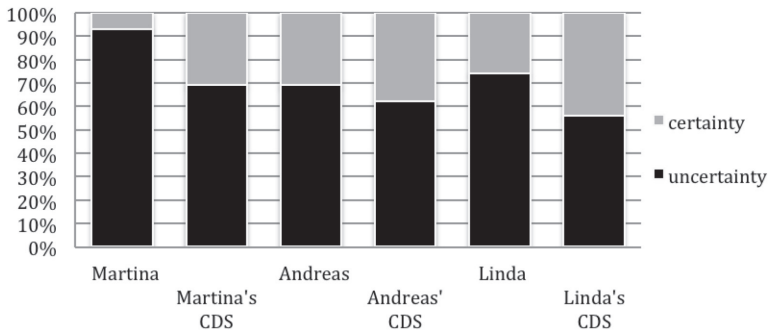


**Figure 2.** Distribution of EM (tokens) in EST
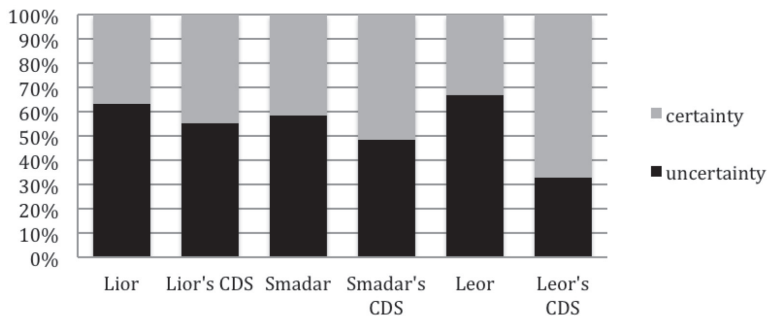
**HEB: Certainty *vs*. Uncertainty**



**Figure 3.** Distribution of EM (tokens) in HEB

## 5.2. EMERGENCE OF EM AND THEIR FEATURES

The children under observation start using EM at age 1;7–2;4 (Table 4 below). Their MLU showing so-called linguistic age at this moment is different. Thus, EM occur at the stage of three-word utterances in RUS and EST, whereas in HEB, standard deviation is too high: it shows the different MLU levels when children start to use epistemic marking.

**Table 4.** First occurrences of (un)certainty markers

| Language | RUS | EST | HEB |
|---|---|---|---|
| Age | 2;1–2;4 | 1;10–2;0 | 1;7–1;11 |
| M (age in months) | 25.3±0.6 | 23±1 | 21.7±2.3 |
| MLU | 2.6±0.5 | 3.13±1.4 | 1.5±2.6 |

### 5.2.1. FIRST EM IN RUS CS

The first cases of EM usage in RUS CS are infrequent for both epistemic domains (Table 5 below).

**Table 5.** Emergence of first markers of (un)certainty in RUS

| | Uncertainty | | | Certainty | | |
|---|---|---|---|---|---|---|
| | Age | EM[13] | MLU | Age | EM | MLU |
| Filipp | 2;1 | 1/3 | 3.3 | 2;5 | 1/1 | 2.4 |
| Liza | 2;2 | 2/4 | 3.1 | 3;10 | 3/7 | 4.6 |
| Kirill | 2;4 | 2/3 | 1.5 | 2;4 | 1/1 | 2.7 |

Qualitative analysis of initial EM shows that children begin with markers mostly expressing a high degree of epistemic semantics of uncertainty: *navernoe* and its colloquial version *naverno* 'probably'.

(1)  2;1  Filipp:  ***naverno**, papa* (podaril igrušku) 'Probably daddy (gave me a toy)'

---

[13]  Here and further (Tables 6, 7) numbers given are in lemmas and tokens.

(2)  2;4  Mother:  *a čto, on* (o magazine) *vnizu?* 'Is it (a shop) at the bottom?'

Kirill:    ***navernoe*** 'Probably'

They are used in statements mainly describing the situation 'here and now' (2) or the near past (1). The focus of epistemic evaluation is on some elements of proposition: particularly, physical activities of 'third persons' (non-participants in a dialogue) or location of objects. The first EMs are used by children in primary functions.

The other uncertainty markers are *po-moemu* 'in my opinion' (3) and *možet* 'may', which is a colloquial variant 'maybe' (4). They occur in the speech of Liza and Kirill within the same month:

(3)  2;2  Liza:    *ne ta* (o sčitalke)*,* ***po-moemu*** 'Not that (about a rhyme), in my opinion'

(4)  2;4  Kirill:  ***možet***, *drugoe*? 'May(be) another?'

The first marker of certainty emerges slightly later in the speech of both boys (5) and much later in the girl's speech (6), but already with some diversity (3 lemmas):

(5)  2;5  Filipp:  *bol'šaja,* ***konečno*** (o pugovice) 'Big of course' (about a button)

(6)  3;10 Liza:    *Mm, potom, da,* ***konečno*** (ja pomoju ruki) 'Mm, then yes, of course (I will wash my hands)'

*Konečno* 'of course' expresses a high degree of confidence (i.e. fulfils the basic function) and is used by the children at the end of their utterances which are declaratives. The children epistemically evaluated the size of objects (5), their own wishes and future actions (6). Importantly, from the beginning, one of them (Kirill) uses EM of both domains as a reply (2) and as part of a question (4) in a dialogue.

### 5.2.2. FIRST EM IN EST CS

The first cases of EM usage in EST CS are infrequent (Table 6 below).

**Table 6.** Emergence of first markers of (un)certainty in EST

|  | Uncertainty | | | Certainty | | |
|---|---|---|---|---|---|---|
|  | Age | EM | MLU | Age | EM | MLU |
| Andreas | 2;0 | 1/3 | 2.5 | 2;3 | 1/1 | 3.4 |
| Martina | 1;11 | 2/5 | 4.7 | 1;6 | 1/1 | 2.7 |
| Linda | 1;5 | 1/1 | 1.7 | 1;5 | 1/1 | 1.7 |

Two children, Linda (1;5) and Martina (1;6), first used the marker of certainty *tõesti* 'really' but it was used as an imitation by Linda and only once by Martina. Because of this randomness and imitative character of the usage, it cannot be considered to have been acquired first. The first marker of uncertainty *vist* 'probably' appears in the speech of Linda at 1;10, Martina at 1;11, and Andreas at 2;0:

(7)  1;11 Mother:  *on* (valmis) *või*? 'Is ready?'
          Martina:  *valmis küll, peab **vist** puhuma* 'Ready yes, must
                    probably blow (to cool down)'

(8)  2;0 Andreas:  *mis sa noonistad* [joonistad]? 'What are you
                    drawing?'
          Father:  *jõehobu* 'A hippo'
          Andreas: *liblika **vist*** 'Butterfly probably'

This marker was used by the children to refer to some kind of activity of themselves or somebody else.

EM of certainty started to emerge a bit later: *tegelikult* 'really' emerged at 2;3 in Andreas' speech (1/1 as an imitation) and *tõesti* 'really, truly' at 2;5 (1/1):

(9)  2;3 Mother:  *aga tegelikult?* 'But really?'
          Andreas: ***tegelikult** ka ei olnud* (midagi kõrvas) '(It) was
                    reallynothing (in my ear)'

(10) 2;5 Andreas: *meil silmad liiguvad, aga meil muumil ei liigu*
                    'Our eyes are moving but the eyes of our Moomin
                    do not move'
          Andreas: ***tõesti** ei liigu* 'Really/truly do not move'

Estonian children started to express epistemic modality with EM of middle degree of uncertainty mostly in non-interrogatives (7, 8), the first interrogatives were used after the age 3;1. The first markers fulfilled the basic function: the children epistemically evaluate the action (7, 10), nature (8) or location of objects (9).

### 5.2.3. FIRST EM IN HEB CS

All the children appeared to start out by expressing uncertainty and only a few months later showed evidence of using certainty markers (Table 7 below).

**Table 7.** Emergence of first markers of (un)certainty in HEB

|  | Uncertainty | | | Certainty | | |
|---|---|---|---|---|---|---|
|  | Age | EM | MLU | Age | EM | MLU |
| Lior | 1;11 | 1/9 | 2.07 | 2;7 | 1/1 | 3.01 |
| Smadar | 1;7 | 1/4 | 2.04 | 1;10 | 1/3 | 2.85 |
| Leor | 1;11 | 1/7 | 2.64 | – | – | – |

The first marker of uncertainty is the modal adverb *ulay* 'maybe' (Smadar at 1;7; Lior and Leor at 1;11). The children use it to make speculations and guesses about positioning (11), the identity of objects (12), and object location (13):

(11) 1;7  Smadar: *ulay kaxa, kaxa ulay, hine ulay kaxa* 'Maybe thus, maybe thus, here maybe thus (=like this)'
(12) 1;11 Lior: *ulay magevet* 'Maybe (a) towel'
(13) 1;11 Leor: *roce od eyfo od, eyfo ulay kan, ulay kan hine ulay sham...* 'Want more where (is) more, where maybe here, maybe here, look maybe there...'

Certainty is first expressed with the adverb *betax* 'sure'. It emerges at a different age for each child (Smadar at 1;10; Lior at 2;7; Leor – none), and is initially used to weaken certainty rather than to reinforce it as in (14, 15).

(14) 1;10 Smadar:  *hu **betax** yaskim lesaxek iti ba-kadur* 'He will
                              surely agree to play ball with me'
(15) 2;7  Lior:     *mi axal? Atem **betax** atem* 'Who ate? You surely
                              you'

The early EM used by HEB-speaking children express middle uncertainty (*ulay* 'maybe') and high certainty (*betax* 'sure'). Both adverbs are most often used in statements to express the girls' speculations.

Thus we have found that the EST and HEB children start to use EM earlier than the RUS ones, specifically before 2;0. However, one could not say that MLU indexes (showing the language age) are very different; bigger differences emerge not between the languages in question, but between the children. In all languages the number of initial epistemic lemmas is not large (1–2); the individual differences in their frequency are more numerous (1–8/9). The main difference between initial EM is their degree (for both (un)certainty), which is high in RUS and middle in EST. In HEB it is middle for uncertainty, high for certainty.

## 5.3. DEVELOPMENT OF EPISTEMIC MODALITY

The children in the study used more uncertainty markers than certainty markers not only in the beginning but also during the whole observation period in all three languages (Figures 4–6). After the emergence of the first certainty markers the children started to use them more often, but they were still not as frequent as uncertainty markers in CS.

### 5.3.1. DIVERSITY AND GROWTH OF THE REPERTOIRE OF EM

The lexical diversity of EM in RUS CS is somewhat smaller than in CDS (on average 7 vs. 10 lemmas, see Table 8 below). Furthermore, the usage made of each marker is between 2–8 times as high in CDS as in CS, a characteristic which may lead to further entrenchment of EM in CS.
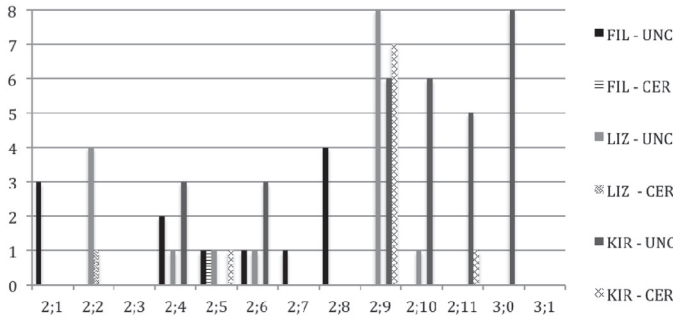
**Figure 4.** (Un)certainty markers (tokens) in the speech of RUS-speaking children
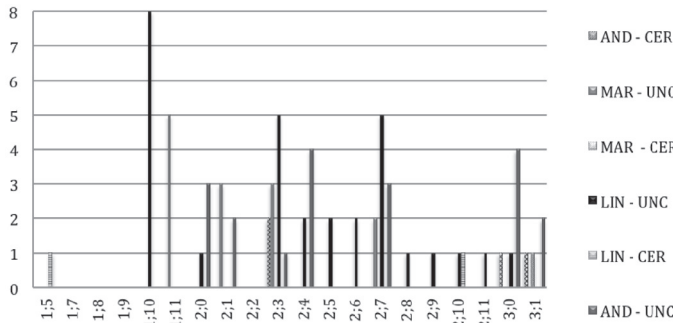


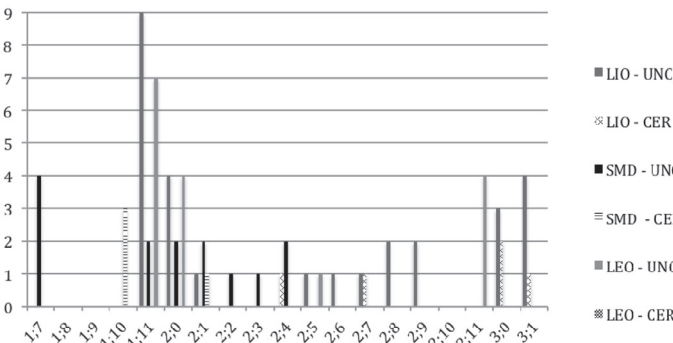**Figure 5.** (Un)certainty markers (tokens) in the speech of EST-speaking children[14]



**Figure 6.** (Un)certainty markers (tokens) in the speech of HEB-speaking children

---

14    For better comparability only one hour of speech is presented here from age 3;0 and 3;1.

**Table 8.** Diversity of the repertoire and frequency of EM in RUS

|  | CS | | CDS | |
|---|---|---|---|---|
|  | **lemmas** | **tokens** | **lemmas** | **tokens** |
| Filipp | 4 | 13 | 7 | 61 |
| Liza | 7 | 34 | 9 | 291 |
| Kirill | 9 | 196 | 13 | 337 |

It is also observed that the children continue to use certain EM constantly after they have emerged so that EM which emerge earlier in CS tend to be used more frequently. Particularly in the domain of uncertainty (Fig. 1), all children start with *navernoe* 'probably'. *Po-moemu* 'in my opinion', *možet* lit. 'may', *možet byt'* 'maybe', *kažetsja* 'it seems', *vidimo* 'apparently' and *vozmožno* 'perhaps' emerge subsequently (two out of three children start with 2 EM of uncertainty). Their distribution in CS shows that *možet byt'* and *možet* 'maybe' rank second after *navernoe* 'probably', with *po-moemu* 'in my opinion' being the least frequent up to 4;0. Acquisition of certainty EM begins with *konečno* 'of course' followed by *dejstvitel'no* 'really', *na samom dele* 'actually' and *pravda* 'truly' lit. 'truth'.

The diversity of EM in EST CS is also not large: uncertainty is represented with 7 EM and certainty with 3. On average, there are 6 in CS and 10 in CDS (Table 9 below).

**Table 9.** Diversity of the repertoire and frequency of EM in EST

|  | CS | | CDS | |
|---|---|---|---|---|
|  | **Lemmas** | **Tokens** | **Lemmas** | **Tokens** |
| Andreas | 7 | 51 | 10 | 380 |
| Martina | 4 | 14 | 9 | 157 |
| Linda | 7 | 102 | 11 | 354 |

The order of emergence of EM in EST CS is the following: *vist* 'probably' (1;10–2;0) > *äkki* 'maybe (coll. speech)' (1;10–2;3) > *võib-olla* 'maybe' (1;11–2;6) > *tegelikult* 'actually' (2;3–2;10) > *muidugi* 'of course' (3;0–4;4) > *ilmselt* 'apparently' (3;0) > *minu meelest* 'in my opinion' (3;0) > *kindlasti* 'for sure' (3;11).

The order of emergence of uncertainty EM was similar in the speech of Linda and Andreas; their next marker was *äkki* 'maybe':

(15)  1;10  Mother:  *mina ei taha neid komme* 'I don't want these candies'

          Linda:    *taha* [tahad] *teda* [seda] ***äkki***? 'Want these maybe?'

The uncertainty marker *võib-olla* 'maybe' emerged next in the speech of all the children: Martina (1;11), Andreas (2;0), and Linda (2;6):

(16)  1;11  Mother:  *jaa sajab lund* 'Yes it's snowing'

          Martina:  ***vuibolla*** [võibolla] *ei sajagi* '*Maybe* it does not snow (running to the window to check)'

This marker refers mostly to the possibility of the situation and the hope that something will happen. *Äkki* 'maybe' was the fourth one to emerge in Martina's speech and was also used in this context:

(17)  2;1  Mother:  *ma usun, et ju ta täna ikka tuleb see linnukene siia* 'I believe that the birdie will come here today'

          Martina:  ***äkki*** *ta tuleb sööma ka* 'Maybe it will come and eat also'

EM that emerge earlier in CS tend to be used more frequently: *vist* 'probably' is the most frequent marker in CS; the most infrequent EM *ilmselt* 'apparently' and *tõenäoliselt* 'probably' emerge later.

In HEB on average 3 EM in CS and 9 in CDS have been used (Table 10 below).

**Table 10.** Diversity of the repertoire and frequency of EM in HEB

|          | CS      |        | CDS     |        |
|----------|---------|--------|---------|--------|
|          | **Lemmas** | **Tokens** | **Lemmas** | **Tokens** |
| Smadar   | 3       | 18     | 3       | 29     |
| Lior     | 5       | 33     | 10      | 147    |
| Leor     | 2       | 16     | 6       | 52     |

The children start epistemic marking with a single marker of uncertainty *ulay* 'maybe'. From 2;0 onward they expand their repertoire of

EM and there is increasing frequency of EM usage. Adverbs marking certainty appear several months later, starting with *betax* 'sure', which, as noted, initially serves to mark attenuated certainty. In addition to the first uncertainty marker, Lior (2;9) starts using the modal verb *yaxol* 'can, be able to' with the copula *be* and the negation word *not – loh yaxol lihyot* 'can't be', and Leor uses (2;11) *kanir'e* 'probably' to mark uncertainty. At 2;1 Smadar uses *batuax* 'certain', and at 3;0 Lior also uses *yaxol lihyot* 'could be' and *carix lihyot* 'must be' to mark certainty. Thus the children share the early adverbial EM but differ with respect to use of other EM and, in particular, use of modal verbs to mark epistemic modality. Compared with RUS and EST, they use a small number of EM.

The present authors have thus managed to uncover some common sequences in the development of (un)certainty, which, in turn, partially coincide in the three languages. Moreover, one can state that the sequence of EM and their frequency in CS is connected. The first EM become the most frequent in CS. Within the domain of uncertainty this means *navernoe* 'probably' (RUS), *vist* 'maybe' (EST), and *ulay* 'maybe' (HEB). Within certainty they are *konečno* 'of course' (RUS), *tegelikult* 'actually, indeed' (EST), and *betax* 'sure' (HEB). Interestingly, in the last case, the EST children begin with this marker, whereas it was the last in our RUS data. A prominent characteristic is that within uncertainty the colloquial versions of EM emerge earlier in all languages.

### 5.3.2. FUNCTIONS OF EM

All RUS and EST children in the early stages use EM in their basic functions, i.e. for expressing their subjective attitude to the proposition or its part. Some cases of EM in secondary functions are present in Kirill's speech after 4;0 (18, where a marker is used to construct a hypothesis) and in Linda's data at 2;0 (19, when she wants to change an unpleasant topic of conversation and introduces a new one with *äkki* 'maybe'):

(18) 5;10 Mother:   *vot eta štuka mne nravitsja, ja ne znaju, začem*
                    *ona nužna* 'I like this thing, I do not know why it
                    is needed'

          Kirill:   **navernoe**, *čtoby ne podorvat'sja srazu* 'Probably,
                    not to explode (himself) right away'

(19) 2;0  Mother:   *vaata mulle nüüd otsa.* 'Look at me now' (Mother
                    had said that the child has to go to eat every meal-
                    time in kindergarten)
                    *kas on kokku lepitud*? 'Do we have a deal?'

          Linda:   *ja.* 'Yes'
                    **äkki** *mängime kulli* 'Maybe we'll play tag'

There are not enough instances to detect a true developmental change in function in HEB. Still, *ulay* 'maybe' is initially used for speculating about objects and their position and only later for making polite requests. *Ulay* 'maybe' is used in CDS both to express caregivers' hypotheses and speculations regarding particular situations or activities, and to make polite requests of the children.

### 5.3.3. SEMANTIC AND PRAGMATIC CHARACTERISTICS OF EPISTEMICALLY MARKED UTTERANCES

Epistemic marking shows a gradual expansion of the children's reflection in respect of surrounding realities: subjects/agents (20) or objects, their presence/absence (21) and action (22):

(20) 2;2  Liza:      *eto mama,* **naverno,** *nalila* (vodu) 'Mom has
                    probably poured (some water)'
(21) 2;0  Andreas:  *mängutoas* **vist** (on klaver) '(The piano is)
                    probably in the playing room'
(22) 2;1  Smadar:   *ze* **ulay** *mafria le-ze* 'It probably blocks this'

Thus, initially, epistemic marking emerges in third-person utterances (i.e. *he/she/it*-utterances) mostly describing the location and identity of objects (see also examples 4–16 above). The epistemic

evaluation of objective situations and/or its components takes place in 70% of cases of epistemically marked utterances in RUS and EST.

From approximately the end of the third year of life, epistemic marking is used in children's utterances about themselves: their own actions (23, 24), including mental ones and inner (psychological) states (25):

(23) 2;10 Linda:   ***tegelt*** [tegelikult] *ma ei pane juustu* 'Actually I will not put cheese (on to the sandwich)'

(24) 2;3  Smadar: ***ulay***, *ani eshal et Pigi eyfo ha-kos shel dubi...?* 'Maybe I'll ask Pigi where the bear's glass is?'

(25) 2;8  Filipp:   ***kažetsja***, *i ja plaču* 'It seems, and I am crying'

After EM were used in first person (*I*)-utterances, i.e. sentences about themselves (20% in RUS, EST), children are able to evaluate the mental or other activity of their partner, i.e. they start to use EM in second-person (*you*)-utterances:

(26) 3;0  Andreas: *sa oled* ***vist*** *väga haige laps* 'You are probably a very sick child'

In HEB *you* is not referred to after *I* – the dichotomy seems to be between subjectless sentences or sentences with 'it' and ones with personal pronouns like *I* and *you*.

(27) 2;3  Smadar: ***ulay*** *ata yaxol lexakot?* 'Maybe you can wait?'

Thus *you*-utterances are marked epistemically (10% in RUS, EST). In Kazakovskaya and Argus (2016) and Kazakovskaya (2017, *forthcoming*) it was shown that the focus of initial epistemic evaluation connected with the child's theory of mind, since it reflects the child's point of view and his/her representations about 'the other'. So children start to use EM, initially for describing some kind of general situation and after that they start to epistemically evaluate the situations connected with themselves and their interlocutors using mental (RUS) and/or modal verbs (RUS, EST).

Characterizing the main pragmatic features of epistemically marked utterances, one can state that in RUS and EST, epistemic marking starts mainly from non-interrogative utterances, which are declaratives and/or answers (20, 21). Later (usually after 3;0) it becomes possible within interrogatives (28, 29). HEB *ulay* 'maybe' is more often used in questions than *betax* 'sure', which is a certainty marker. HEB-speaking children ask questions early using *ulay* 'maybe', although up to age 1;11–2;0 they tend to use it in statements rather than questions.

(28) 3;6  Liza:      *a **možet**, pojdjom pokataemsja na gorke sejčas?* 'Or may(be) go to ride on the hill now?'

(29) 3;1  Andreas:  ***tegelikult** sina, kas sa tahad mulle nii palju anda?* 'Actually you, do you want to give me so much?'

(30) 2;2  Smadar   ***ulay** tash'ir oto ba-sal?* 'Maybe you (can) leave it in the basket?'

This sequence repeats the general order of the acquisition of non-interrogatives before interrogatives by children: this was demonstrated for RUS in Kazakovskaya (2011). Although interrogatives are quite frequent in everyday speech, they are considered as more cognitively-complicated utterances. At the same time, epistemically marked questions emerge late developmentally due to the cognitive complexity of a situation which involves the evaluation of interlocutor's attitudes. It also might be culturally important because it would not be common or even polite to evaluate something connected with your interlocutor's action or mental state.

## 5.4. INPUT-OUTPUT RELATIONSHIPS

Analysis of the relationships between EM in CS and CDS has revealed two correlations. The first of them lies in the sphere of frequency. There are many more EM in CDS than in CS, and a child whose parent uses more EM also has more markers in their own

speech. The second correlation concerns the diversity of EM (Tables 8–10 above). These similarities suggest that parents' and children's use of EM go hand in hand. To illustrate this, we have carried out a developmental analysis of EM for each dyad (see, for instance, Figures 7–9 below).
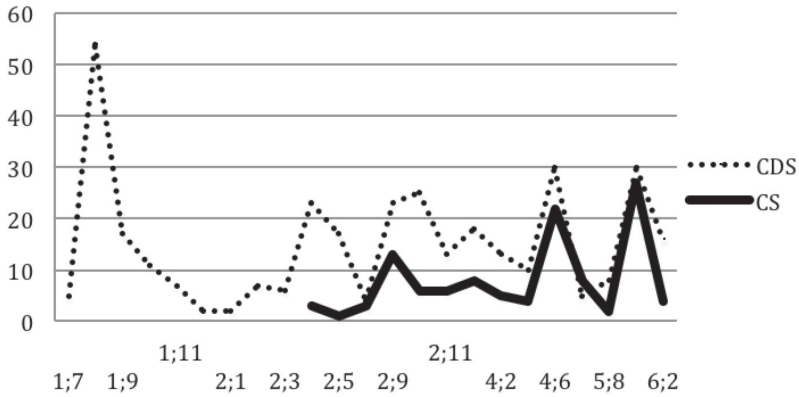
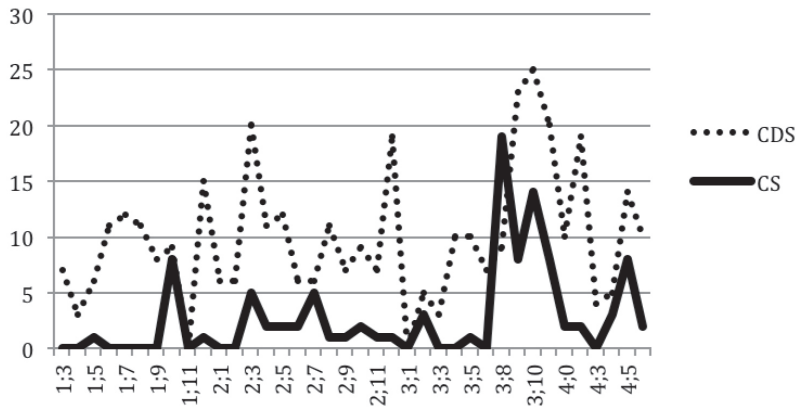Figure 7. Usage of EM by RUS child Kirill and his caregiver (tokens)

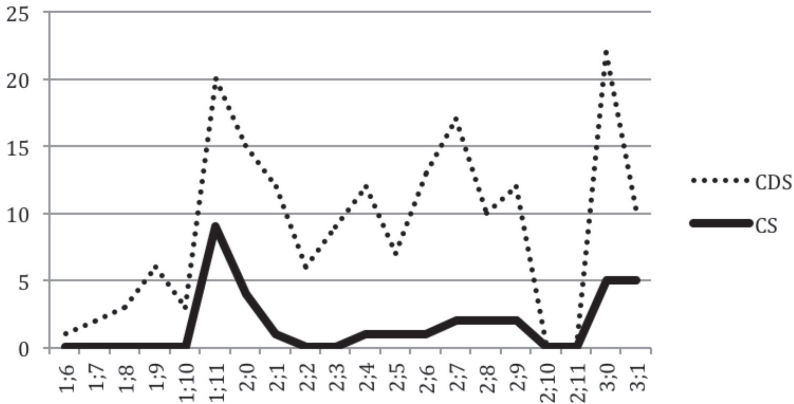Figure 8. Usage of EM by EST child Linda and her caregiver (tokens)

**Figure 9.** Usage of EM by HEB child Lior and her caregiver (tokens)

Comparison of developmental trends in CDS and CS shows that most children have some kind of 'epistemic peaks'. Particularly, in the data of RUS Kirill (at 2;9, 4;6, 5;10, see Fig. 4), EST Linda (at 3;8, see Fig. 5) and HEB Lior (at 1;11, 3;0, see Fig. 6), these are quite prominent. Although such peaks, which are found in the other dyads under observation, may be taken as evidence for fine-tuning of CDS to the children's 'epistemic development', it must be pointed out that the children use EM either in response to their presence in the preceding utterance of the caregiver or in a wider conversational context. Another factor affecting the use of EM in CS is the nature of 'adult–child' interaction and the ongoing activity (Kazakovskaya, *forthcoming*, Kazakovskaya, Argus 2016).

## Conclusions

Developmental analysis of EM of (un)certainty in the languages under observation has supported the order of emergence founded in Kazakovskaya (2011, 2016) and Kazakovskaya and Argus (2016): uncertainty before certainty. The main reason why children start to mark uncertainty can be the higher frequency of uncertainty EM in CDS.

EM occur at different ages (in EST and HEB this process starts slightly earlier than in RUS) but does so at the multi-word utterances stage. The difference between RUS and EST as concerns MLU (in words) can be explained by the typological nature of these languages. In EST, grammatical information is usually attached to the stem (a separate suffix for each piece of information), whereas in RUS several grammar meanings are contained within a single suffix, and there is more use of separate words (e.g. prepositions) in RUS compared to EST.

The input-output relationship reveals the influence of frequency, diversity and fine-tuning of CDS. Thus the sequence of EM emergence in CS coincides with their frequency in CDS. The lack of striking differences in the acquisition of certainty and uncertainty allows us to conclude that typological differences are less significant for this process, in contrast to input frequency and diversity of EM along with their language-system prototypicality: e.g. in EST high-degree uncertainty EM are not frequent in CDS, which may be one of the reasons for their late emergence in CS.

In all three languages under investigation the children's epistemic marking begins within non-interrogatives (declaratives and/or declarative answers within question-reply units). This is due to the latter's greater simplicity when compared with the cognitive complexity of interrogatives. Children's epistemic marking develops from the evaluation of objective situations/its components as location or guesses about the identity of an object to mental reflection concerning their partner through to sentences about themselves, using mental (RUS) and/or modal (EST, HEB) verbs. In RUS, epistemic marking begins with high-degree of EM; in EST, it is at the middle-degree, whereas HEB has middle uncertainty and high certainty EM. The functions of EM develop from pure epistemic evaluation of the proposition towards more complex functions like conversational strategy (e.g. change of an unpleasant topic) and politeness.

The preliminary sketch of the cultural schema of epistemic marking may be described in the following way: in RUS and EST, EM

are used for expression of uncertainty rather than certainty in both CDS and CS. RUS and EST parents use EM in their dialogues with children with a different frequency compare to 'adult–adult' conversations (for a more detailed description see Kazakovskaya, Argus 2016). Besides this, RUS caregivers tend to use a rather high-degree EM of both domains, whereas EST ones use the middle-degree of uncertainty EM for evaluating some kind of general situation or location of objects rather than other situations, e.g. somebody's action. Moreover, they use epistemically marked sentences for describing the speaker's own actions or mental states rather than the interlocutor's. The directness of HEB can be observed in epistemic marking: the number of epistemically marked utterances was smaller than in RUS and EST and the overall use of certainty markers in HEB CDS as well as CS was relatively higher than in the two other languages studied here..

Thus wealth, frequency, the degree of semantics and pragmatic functions of EM, as reflected in CDS, are better predictors of early emergence and diverse repertoire of EM in CS than general morphological typology. Communicative styles of caregivers and cultural norms are also important for epistemic acquisition. On the basis of our data and preliminary analysis, the general traditions of speaking about the degree of confidence of different propositions (context schema of epistemic marking) and behavior connected to evaluation of, for example, interlocutor's ideas or feelings (emotion schema) vary cross-culturally. On the whole, the picture obtained is like the beginning of a cultural schema for epistemic evaluation in 'adult–child' communication, which needs to be compared to 'adult–adult' dialogues and requires further verification using various language data.

REFERENCES

Argus, Reili, *forthcoming*. Acquisition of requests in Estonian. In Stephany, Ursula; Akcu-Koç, Ayhan (eds.) Development of Modality in First Language Acquisition: A Cross-linguistic Perspective. Berlin: de Gruyter.

Beljaeva, Elena I. 1990. Reliability. In Aleksandr V. Bondarko (ed.), Theory of functional grammar. Temporality. Modality. Leningrad: Nauka, 157–170 (in RUS).

Berman, Ruth A. 2011. Revisiting impersonal constructions in Modern Hebrew: Discourse-based perspectives. In Andrej Malchukov & Anna Siewierska (eds.), Impersonal constructions: A cross-linguistic perspective, 323–356. Amsterdam: John Benjamins.

Berman, Ruth A. 2014. Acquiring and expressing temporality in Hebrew: A T/(M/A) language. – SKASE Journal of Theoretical Linguistics 11(2), 2–29.

Birjulin, Leonid A.; Elena E. Kordi 1990. Main types of modal meanings allocated in linguistic literature. In Aleksandr V. Bondarko (ed.). Leningrad: Nauka, 67–71 (in RUS).

Boye, Kasper 2016. The expression of epistemic modality. In Jan Nuyts and Johan Van Der Auwera (eds.) The Oxford Handbook of Modality and Mood. Oxford: Oxford University Press, 117–140.

Bulygina, Tatjana V.; Aleksej D. Šmelev 1992. Communicative modality. In Arutjunova, Nina D.; Bulygina, Tatjana V. (eds.). The Human Factor in Language: Communication. Modality. Deixis. Moscow: Nauka, 110–137 (in RUS).

Choi, Soonja 2006. Acquisition of Modality. In W. Frawley (ed.) The Expression of Modality. Berlin: Mouton de Gruyter, 141–171.

Coffin Amir, Edna; Shmuel Bolozky 2005. A Reference Grammar of Modern Hebrew. Cambridge: Cambridge University Press.

de Haan, F. 2005. Encoding speaker perspective: Evidentials. In Z. Frajzyngier, A. Hodges & D. S. Rood (eds.), Linguistic diversity and language theories. Amsterdam: John Benjamins. doi: 10.1075/slcs.72.18haa

Dromi, Esther 1980. Modality in Modern Hebrew. In Patricia Hamel; Ronald Schaefer (eds.), Kansas Working Papers in Linguistics 5, 99–115.

Erelt, Mati 2017. Öeldis. Eesti keele süntaks. Eesti keele varamu 3. Tartu: Tartu Ülikooli Kirjastus.

Gvozdev, Aleksandr N. 1949. The formation of the Russian grammatical system in the child. 2 vols. Moscow: Academy of Pedagogical Science of the Russian Federation Press (in RUS).

Hickmann, Maya; Dominique Bassano 2016. Modality and Mood in First Language Acquisition. In Nuyts, Jan; Auwera Johan Van Der (eds.) The

Oxford Handbook of Modality and Mood. Oxford: Oxford University Press, 430–447.

Hoff, Erika 2006. How Social Contexts Support and Shape Language Development. – Developmental Review 26, 55–88.

Jakovleva, Ekaterina S. 1994. Fragments of the Russian language picture of the world (models of space, time and perception). Moscow: Gnozis (in RUS).

Kazakovskaya, Victoria V. 2011. Questions and answers in "adult–child" dialogue. Moscow: Editorial URSS (in RUS and ENG).

Kazakovskaya, Victoria V. 2016. Input–Output Relations in Early Epistemic Modality in Russian: a corpus-based approach. In Li Chen (ed.), Proceedings of the 3rd International Conference on Applied Social Science Research. Paris, Amsterdam, Beijing: Atlantis Press, 423–426.

Kazakovskaya, Victoria V. 2017. Language and cognitive aspects of the acquisition of epistemic modality. – Acta Linguistica Petropolitana 13(3), 542–575 (in RUS).

Kazakovskaya, Victoria V., *forthcoming*. Epistemic modality in Russian child language. In Stephany, Ursula; Akcu-Koç, Ayhan (eds.) Development of Modality in First Language Acquisition: A Cross-linguistic Perspective. Berlin: de Gruyter.

Kazakovskaya, Victoria V.; Reili Argus 2016. Acquisition of epistemic marking in Estonian and Russian. – Estonian Papers in Applied Linguistics 12. 57–80.

Kehayov, Petar 2008. Interactions between grammatical evidentials and lexical markers of epistemicity and evidentiality: a case study of Estonian and Bulgarian. – Wiener Slavisticher Almanach, Sonderband 72, 165–201.

Kiefer, Ferenz 2000. Jelentéselmélet [Semantic theory]. Budapest: Corvina.

Krauze, Marion 2004. Modal markers in child speech: Development of functions and system. In Stella N. Ceitlin (ed.) Children's speech as an object of linguistic research. St. Petersburg: Nauka, 136–139.

Lieven, Elena V. M. 1994. Crosslinguistic and cross-cultural aspects of language addressed to children. In: Gallaway, Clare; Richards, Brian J. (eds.) Input and interaction in language. Cambridge: Cambridge University Press, 56–73.

MacWhinney, Brain 2000. The CHILDES project. Tools for analyzing Talk. Electronic edition. Part 1: The CHAT transcription format.

Matsui, Tomoko 2014. Children's understanding of linguistic expressions of certainty and evidentiality. In Danielle Matthews (ed.), Pragmatic Development in First Language Acquisition. Amsterdam: John Benjamins, 295–316.

Müürsepp, Mari-Liis 2015. Katselised meetodid modaaladverbide evidentsiaalsuse ja episteemilise modaalsuse uurimisel. MA-thesis in Tartu University.

Nishida, Heiko 1999. A cognitive approach to intercultural communication based on schema theory. – International Journal of Intercultural Relations 23(5), 753–777.

Noveck, Ira; Simon Ho; Maria Sera 1996. Children's Understanding of Epistemic Modals. – Journal of Child Language 23(3), 621–643.

Nuyts, Jan 2016. Epistemic Modality, Language, and Conceptualization: A Cognitive-pragmatic Perspective. John Benjamins Publishing Company.

Ochs, Elinor; Bambi B. Schieffelin, (eds.). 1986. Language Socialization Across Cultures. Cambridge: Cambridge University Press.

Oficerova, Ekaterina A. 2005. The expression of modal meanings of possibility and necessity in Russian child speech. St. Petersburg: Institute for Linguistic-Studies of the Russian Academy of Sciences dissertation.

Ovčinnikova, Irina G.; Inna A. Uglanova; Marion Krauze 1999. On estimation of the certainty/uncertainty degree of utterances by children of two age groups. In Stella N. Ceitlin (ed.) Problems of ontolinguistics 1999. St. Petersburg: Hertsen University Press, 132–133 (in RUS).

Pajupuu, Hille 1995. Cultural context, dialogue, time. Tallinn: Estonian Academy of Sciences. Institute of the Estonian Language. https://www.eki.ee/teemad/kultuur/context/context.html (28.09.2018)

Palmer, Frank R. 1986. Mood and modality. Cambridge: Cambridge University Press.

Panfilov, Vladimir Z. 1977. The role of modality in the construction of sentence and judgment. – Voprosy jazykoznanija 4, 37–48 (in RUS).

Sedov, Konstantin F. 2004. Discourse and personality: the evolution of communicative competence. Moscow: Labirint (in RUS).

Russkaja grammatika/ Russian Grammar. 2005 (1980). Vol. 2. Moscow: Nauka (in RUS).

Stephany, Ursula; Ayhan Akcu-Koç (eds.), *forthcoming*. Development of Modality in First Language Acquisition: A Cross-linguistic Perspective. Berlin: de Gruyter.

Stoljarova, Irina V. 1992. Expressing subjective and modal relations in child speech. In Stella N. Ceitlin (ed.), Child language: Linguistic aspect. St. Petersburg: Obrazovanie, 86–95 (in RUS).

Sharifian, Farzad 2017. Cultural Linguistics. John Benjamins Publishing Company.

Tüüts, Laura; Reili Argus 2016. Episteemilise modaalsuse leksikaalsete väljendusvahendite tajumisest: *arvatavasti*, *äkki* ja *võib-olla*. – Eesti ja soome-ugri keeleteaduse ajakiri 7(2), 187–208.

Uziel-Karl, Sigal, *forthcoming*. Modality in child Hebrew. In Stephany, Ursula; Akcu-Koç, Ayhan (eds.), *forthcoming*. Development of Modality in First Language Acquisition: A Cross-linguistic Perspective. Berlin: de Gruyter.

Valdmets, Annika; Külli Habicht 2013. Episteemilistest modaalpartiklitest eesti kirjakeeles. Eesti ja soome-ugri keeleteaduse ajakiri [Journal of Estonian and Finno-Ugric Linguistics], 4 (1), 205–222 (in EST).

Vinogradov, Viktor V. 1947. Modal words and particles. In Vinogradov, Viktor V. Grammatical doctrine of the word. Moscow, Leningrad: Učpedgiz, 725–744 (in RUS).

Wiemer, Björn 1992. Acquisition of modal meanings during ontogenesis (based on Russian and English). In Stella N. Ceitlin (ed.), Child language: Linguistic aspect. St. Petersburg: Obrazovanie, 131–144 (in RUS).

RESÜMEE

## EPISTEEMILISE MODAALSUSE VARANE OMANDAMINE TÜPOLOOGILISELT ERINEVATES KEELTES: EESTI, VENE JA HEEBREA KEELES

Artiklis on käsitletud episteemilise modaalsuse, täpsemalt tõenäosuse ja tõsikindluse väljendamise vahendite arengut varasel keeleomandamise perioodil. Vaatluse all on kolm tüpoloogiliselt ja kultuuriliselt erinevat keelt – vene, eesti ja heebrea keel. Analüüsimaterjali moodustavad üheksa ükskeelse lapse (kolm last iga keele kohta) spontaanse kõne lindistused vanuses 1;3 kuni 6;2 eluaastat. Tõenäosuse ja tõsikindluse väljendusvahendite analüüs hõlmab nii laste kui ka lastele suunatud kõnet.

Kõik vaatlusalused lapsed alustavad tõenäosuse ja tõsikindluse väljendamist teise eluaasta lõpus. Esialgu kasutavad lapsed episteemilise modaalsuse markereid, et anda tõenäosushinnang mingile situatsioonile üldisemalt, seejärel laieneb kasutus mentaalsete seisundite tõenäosuse hindamisele. Kõige hiljem hakkavad lapsed kasutama episteemilise modaalsuse markereid sekundaarses funktsioonis ehk mitte niivõrd tõenäosuse kui pigem viisakuse väljendamiseks või ebamugava vestlusteema muutmiseks. Tõenäosuse väljendamine omandatakse kõigis kolmes keeles varem kui tõsikindluse väljendamine, ka on tõenäosusmarkereid materjalis rohkem kui tõsikindluse markereid ning seda kogu vaatlusperioodi jooksul. Võrreldes laste enda ja lastele suunatud kõne andmeid, ilmneb, et omandamiskäiku mõjutab nii markerite üldine sagedus täiskasvanukeeles kui ka see, millise astme (nt kas vähe, keskmiselt või väga tugevat tõenäosust või tõsikindlust väljendavaid) markereid vanema kõne lapsele kõige enam ette annab. Näiteks kui vene vanemate kõnes on lapse varases vanuses ainult väga tugevat tõenäosust ja tõsikindlust väljendavad markerid, siis eesti vanemad eelistavad

keskmist tõenäosust väljendavaid ning ei kasuta tõsikindlust väljendavaid markereid alguses peaaegu üldse.

Kolme keele erinevus ilmneb kõige ilmekamalt markerite kasutussageduses: heebrea kui kultuuriliselt väga otsese suhtlusstiiliga keelt kõnelevad vanemad kasutavad tõenäosust väljendavaid markereid väga vähe, kõige enam leidub neid markereid vene keelt kõnelevate vanemate ja laste kõnes.

**Võtmesõnad:** esimese keele omandamine, lapsele suunatud kõne, tõenäosus, tõsikindlus