Pierre-Yves Modicom

8 Epistemic verbs in Scandinavian languages: Person-bound paths of functional specialization

Abstract: The aim of this chapter is to examine (i) if and how some properties deemed typical of epistemic verbs in first-person contexts carry over to second-person contexts, and (ii) if different paths of pragmatic specialization can be identified for first- and second-person epistemic verbs. This contribution is intended as a pilot study, based on a set of opinion verbs in three neighbouring languages: Danish, Norwegian and Swedish. Using frequency data from the TenTen familiy of corpora, I show that there is indeed a massive flip from almost exclusively declarative contexts in the first person singular towards a large share of interrogative contexts in the second person, which is consistent with previous findings on the "interrogative flip" in the typological literature. Many other features of epistemic verbs, including their much-discussed "parentheticality", seem to be lexically-driven, with each verb displaying its own statistical preferences and person not being a major factor in variation. In all three languages, the cognates of English *think*, which display a larger range of non-experiential uses than other epistemic verbs, also behave differently with respect to grammatical person. Their deictic-epistemic properties are less significant. On the other hand, these verbs also give rise to the most solid examples of pragmaticalization in the second person, in a specific slot, the utterance-final position, which corroborates a secondary expectation of the study, as to the topological specialization of subjective, first-person epistemic verbs vs intersubjective-second person epistemic verbs.

Keywords: Egophoricity, epistemic verbs, parentheticals, person, Scandinavian

8.1 Introduction

In typological research on epistemicity in interaction, the concept of egophoricity is used to designate the grammatical marking of asymmetric epistemic authority between speech act participants.

Acknowledgements: I would like to thank the reviewers and the editors for their feedback, which led to considerable changes for the better. All remaining mistakes and inaccuracies are mine.

Pierre-Yves Modicom, Université Lyon 3

https://doi.org/10.1515/9783111516233-008

Egophoricity is closely connected with evidentiality, the grammatical marking of access to knowledge. Evidentiality relates primarily, though not exclusively, to the source of knowledge. Egophoricity conveys speaker involvement (see the introduction to the volume). Grammatical egophoricity is commonly observed on certain sets of verbs, most prominently verbs of cognition and emotion, but also, to some extent, verbs of volition and of intention (San Roque, Floyd and Northcliffe 2018). Overall, psychological verbs are the locus of egophoric marking. A reasonable explanation for this is the fact that these verbs are strongly experiential and express states of affairs to which the experiencer has immediate privileged access (see Hargreaves 2018 from a typological point of view). In other words: speaker involvement, the key factor for egophoric marking, overlaps with direct access to the state of affairs, a crucial matter for evidential strategies.

In declarative contexts, the speaker enjoys a specific epistemic authority when they talk about their own mental states (see Davidson 1984; but also Russell 1910). This semantic feature connects epistemic verbs to the realm of deixis and indexicality: the semantic mechanisms behind first-person statements about one's mental states are not representational or symbolic, they correspond to a case of direct experiential access to what is meant. From that perspective, egophoricity is a grammatical manifestation of the intrinsically deictic nature of epistemic meaning (see for instance Mitchell 2009; Hassler 2010; Abraham and Leiss 2012; Leiss 2012).

If we assume epistemic strategies to be sensitive to deixis, we can expect them to react to the interaction of speech act types and grammatical person. In many languages, egophoric features associated with first-person psychological predicates in a declarative context tend to carry over to the second person within an interrogative context. This phenomenon has been identified under the name "interrogative flip" in research on Southern American languages (see San Roque, Floyd and Northcliffe 2017:128; San Roque, Floyd and Northcliffe, Bergqvist and Kittilä 2019:10; Bergqvist 2021; Bergqvist and Grzech 2023; Faller 2024 among many other). In these egophoricity-marking languages, grammatical markers on the verb referring to the speaker in assertions refer to the addressee in questions.

The main goal of this chapter is to investigate to what extent this typically egophoric phenomenon can be observed in Scandinavian languages, which do not display egophoricity as a salient grammatical category – thus corroborating the hypothesis that the functional mechanisms behind grammatical egophoricity are grounded on a pragmatic (in this case: deictic) base. To distinguish possible lexical patterns of co-occurrence from the grammatical category of egophoricity, the more general functional bias will be labelled "deictic-epistemic". If the assumption is correct, deictic-epistemic biases manifest the overarching deicticity of epistemic markers. Proper egophorics are deictic grammatical forms, much akin to Jakobson's "shifters", for example (Jakobson 1971 [1957]). But above all, deixis is a general

mechanism of language, by which symbolic and indexical material are combined to entrench representational contents into the context of production and reception. In that sense, using lexical and grammatical material to flag the epistemic authority of the speech act participants in the context of production is a genuinely deictic process.

To assess the similitude between deictic-epistemic biases and the categorial mechanisms of egophoricity, this chapter offers an empirical study relying on data retrieved from large corpora of Danish, Swedish and the Bokmål written variety of Norwegian.

The rest of this chapter is constructed as follows: Section 8.2 presents the state of the art; Section 8.3 presents the methodology of the study. Section 8.4 is devoted first to illocutionary values, and then to the placement of epistemic sub-clauses in the utterance. The conclusions can be read in Section 8.5.

8.2 State of the art

8.2.1 Epistemic authority and the idiosyncrasies of parenthetical verbs

In the Wittgensteinian tradition of philosophy of mind, and in the discussions borne out of Wittgenstein's and Anscombe's work (Anscombe 1957), it has long been claimed that the type of beliefs about knowledge informing the meaning of verbs like *know* or *believe* or *think* is actually not the same in the first person and in the other persons.[1] Ultimately, the verb *think* does not mean the same process in *I think* and in *she thinks*, because reports about third-person thinking are instances of descriptive knowledge ; on the other hand, self-reports on one's own thoughts do not involve descriptive knowledge of one's thoughts, but direct acquaintance with them. In other words : the meaning of *think*, *believe* or *know* is such that the difference between *I think* and *she thinks* cannot be fully captured by positing a unique meaning for *think* and compositionally combining it with the meaning of *I* vs. *she.* Person-bound epistemic authority implies that in declarative contexts, the first person of thought verbs has something special which is not captured by truth-conditional semantic descriptions of these verbs : *I think* or *I believe* stand in the vicinity of speech verbs in the first person.

1 As early as 1910, Russell's seminal paper on "knowledge by acquaintance" vs "knowledge by description" already took first-person experience of mental states and pain as a point of departure (Russell 1910).

One of the earliest consequent implementations of this idea was made by J.O. Urmson in the 1952 paper in *Mind* where he coined the notion of "parenthetical verbs" for these items – a label which has been used in linguistics and in philosophy ever since. From the point of view of theoretical pragmatics, Anscombe (1957) has insisted upon the fact that first-person epistemic and intentional predicates are cases of non-descriptive (self-)knowledge. She maintains that any first-person statement involving a psychological verb enjoys a specific status. Similarly, many theoricists and philosophers contend that first-person psychological predicates taking a complement clause, just like first-person speech act verbs in the same syntactic contexts, are semantically downgraded to expressive meta-indications about the speaker's attitudes towards the content of the complement clause. The object argument of the epistemic verb is in fact the real descriptive content in the utterance. This functional demotion of the epistemic predicate may be pragmatically motivated, but it can be manifested syntactically by the disruption of the subordination link between the epistemic verb and the content clause. In that case, the epistemic construct is sometimes dubbed a "comment clause" (Heine and Kaltenböck 2021).

8.2.2 Parentheticality and embedding strategies

A survey of the huge body of grammatical literature on psychological verbs would go far beyond the limits of the present paper. Parenthetic properties have to do with the fact that mental verbs take a propositional argument; this propositional argument can be manifested by a full clause.

From a syntactic point of view, the clause denoting a mental content ought to be a subordinate clause (a "completive clause" or "noun clause" in school terminology). However, mental verbs display distinctive formal properties and these subordinate clauses may not be embedded through the same means as other completive clauses (see the discussion on "complement-taking predicates" in Boye and Harder 2021). Among these specific features, the most striking is the possibility to build the sentence without using a complementizer, for instance in example (1):

(1) *Jeg tager det bare som en kompliment, at* du synes *min*
I take it only a a compliment that you think my
side er pæn. (Danish)[2]
site is clean
'I take it only as a compliment that you think my website is clean.'

2 All Danish examples are extracted from the daTenTen20 corpus on SketchEngine, last retrieved on July 23rd, 2024.

Here, *du synes* 'you think' takes a clausal argument without any marker of subordination. This absence of complementizer can be interpreted as syntactic deletion, as phonetic reduction or as a construction *per se*, associated with a specific kind of syntactic government. In Danish, for instance, we can assume this clausal argument to still be syntactically dependent on the epistemic verb, thanks to the existence of examples of the subject – adverb – finite verb order that is restricted to subordinate clauses. In (2), *alvorligt* 'seriously' stands between the subject and the finite verb, as does *ikke* 'not' in (3):

(2) *Hvis ikke, synes jeg du alvorligt skal overveje* at slutte
If not think I you seriously should.PRES consider to end
det her nu. (Danish)
that here now
'If not, I think you should seriously consider ending this now.'

(3) *Personligt synes jeg du* ikke *skulle have brugt omkring 50% af din spalte plads, på et fordrag du syntes var kedeligt.* (Danish)
Personligt synes jeg du ikke skulle have brugt
Personally think I you not should.PAST have used
'Personally, I think you shouldn't have spent about 50% of your column space on a lecture you thought was boring.'

Note, however, that this issue has to be judged separately for each language: in High German, for instance, where there is an OV / V2 asymmetry between V2 main clauses and verb-final (OV) subordinate clauses (as in 4), complementizer-free clausal arguments behave like main clauses example (5):

(4) *Glaubst du, dass das unterhaltend ist ?* (German)
Think you that it entertaining is
'Do you think that it's entertaining?'[3]

(5) *Glaubst du, das könnte für Probleme sorgen ?* (German)
Think you it could PREP problems cause
'Do you think that could cause problems?'

3 German examples are extracted from the deTenTen20 corpus on SketchEngine, last retrieved on July 23rd, 2024.

Blanche-Benveniste (1989) calls the corresponding structure in French "weak government" (*rection faible*). In what follows, I will speak of "weak embedding" for the complementizer-free construction in (2) or (3) and of "strong embedding" for the canonical construction in (6) below, with the subordinator *at*:

(6) *Hvis du synes* at du har hørt melodien før, *så er det nok Blacks "Wonderful Life" fra 1987.* (Danish)
Hvis du synes at du har hørt Melodi-en før
if you think that you have heard melodie-DEF before
'If you think you've heard the melody before, it's probably Black's "Wonderful Life" from 1987.'

Another salient well-known feature of epistemic verbs is the fact that they are can easily be bypassed by *wh*-movement, as in (7):

(7) *Så er det op til dig, hvad* du synes *ser mest interessant ud.* (Danish)
. . . *hvad du synes ser mest interessant ud*
what you think looks most interesting out
'Then it's up to you what you think looks most interesting.'

Furthermore, epistemic verbs in the first person are highly frequent sources of pragmaticalized discourse markers (e.g. Schiffrin 1987; Günthner and Imo 2003; Aijmer 2011 and many other works).

This brings us to parentheticality in the narrow sense, which I define here as the constructional status of a complement-taking predicate whose syntactic dependency relationship with its propositional complement is partly or totally deficient. There is a parenthetical construction at hand when a syntactic cluster organized around the verb displays features of "structural independence" (Dehé and Kavalova 2007:1) from the clausal content over which it takes scope. Complementizer deletion can be regarded as a bridge from standard complementation towards parentheticality proper, which is manifested by the island phenomenon in (7) or by the use of psychological verbs in "comment clauses" and ultimately their pragmaticalization into discourse markers.

Some of these parenthetical properties are also attested for other kinds of verbs which take state of affairs or propositional contents as arguments and are sensitive to the person of the subject, such as speech act verbs. In some languages, they are also well-attested for other persons than the first (see, for instance, Martineau 1993 on Québec French). This raises considerable issues: is parentheticality linked to epistemicity or is it a relevant phenomenon for all verbs denoting attitudes towards postulated states of affairs? And more crucially for any attempt at

connecting parentheticality with egophoricity, are parenthetical constructions sensitive to grammatical person? As it seems, they are not ruled out from second- or third-person contexts, but there could well be either a cross-linguistic accessibility hierarchy, or a quantitative bias towards the first person, or both.

Bearing this in mind, I now turn to the language-specific analysis. I will make use of written data from three Scandinavian languages: Danish, Norwegian (Bokmål) and Swedish. Note that there are already previous studies dealing with these phenomena in oral Swedish, most notably Dahl (2000) and Bergqvist (2021), which tend to corroborate the idea that at least some mental verbs in this language display egophoric behavior. This is manifested by a gap in frequency, with mental verbs tending to occur in the second person in questions, and in the first person in assertions. As we shall see, these claims carry over to Danish and Norwegian, even in a corpus of written discourse.

8.3 Corpus-based investigation

8.3.1 Analytical issues

First, our general problem must be reformulated into one or several testable hypotheses. the following investigation is based on a sample of utterances involving first- or second person epistemic verbs as well as a clausal unit which, from a semantic point of view, designates the mental content taken as argument by the verb. The aim is to determine whether epistemic predicates, especially "parentheticalized" ones, are person-bound deictics replicating some properties of egophorics. If this is the case, we expect the following findings:

– In terms of frequency, the first person singular is strongly associated with declarative illocutionary force, and the second person, with interrogative illocutionary force. That being said, the "illocutionary force" has to be defined as the force of the sub-speech act containing the mental verb: semi-autonomous pragmaticalized units like "comment clauses" can bear an illocutionary value of their own, which needs not coincide with the illocutionary value of the larger utterance. A critical example for this mismatch would be a question tag, which is an interrogative parenthetical micro-unit, but very often acts as a mere modifier of an assertive main clause.
– In these expected canonical patterns, epistemic verbs show stronger signs of parentheticization, starting with complementizer omission, up to pragmaticalization as a discourse marker.

A subsidiary question has to do with the placement of these discourse markers. In the light of recent research, the position of pragmaticalized markers is sensitive to their specialization for subjective vs. intersubjective usage (Beeching and Detges 2015; Van Olmen and Šinkūnienė 2021). Intersubjective markers tend to occur in the "right" (i.e. final) periphery of the utterance or of the turn; subjective markers tend to occur in the "left" (i.e. initial) periphery). The forms undergoing a pragmaticalization process are not the same at the beginning and at the end of an utterance, nor in a declarative and in an interrogative context. Especially, if we are right in assuming a continuum between personal deixis and epistemic deixis, we expect second-person parentheticals to prefer the final position of the clause, with a bias towards interrogative uses.

8.3.2 Data

The following study relies on data from Swedish, Danish, and Norwegian (Bokmål), taken from the TenTen family of corpora, retrieved via SketchEngine. We are not dealing with interactional data taken from oral contexts. A consequence of this bias towards written discourse is that the functions of discourse markers in turn-taking will be left out of consideration here. We can also expect a slight conservative bias and a lesser frequency of completely pragmaticalized occurrences. Conversely, it means that these corpus data are suitable for an analysis of incipient grammatical conventionalization as well as for the basic functions of parenthetical verbs.

Due to the perspective adopted for this chapter, I focus on utterances where the verb occurs together with a clausal unit which, from a semantic point of view, represents the content of thought. I leave aside contexts where the epistemic verb takes an NP (or a pronoun) as its content argument. In the first place, standard embedding uses and parenthetical constructions are all taken into consideration. The constructional distinction is studied in a second step.

The verbs considered for the analysis are the following:

– Danish : *tro* ('believe', 'think'), *tænke* ('think'), *synes* ('it seems' / 'me thinks' / 'I think').
– Norwegian Bokmål : *tro* ('believe', 'think'), *tenke* ('think'), *synes* ('it seems' / 'I think').
– Swedish : *tro* ('believe', 'think'), *tycka* ('think', 'regard'), *tänka* ('think').

Several parameters must be kept in mind for the analysis:

– From a semantic point of view, one should distinguish between experiential vs non-experiential (agentive) kinds of epistemic states (Modicom 2025:324–333).

For instance, English *I believe* has an experiencer subject and opinion verbs are generally experiential verbs: they express mental states that are not controlled by the subject; but at least some of these verbs can display non-experiential, agentive readings. In English, *to think about something* can be used to mean "to reflect about something, to consider something". This point is important because the cognate verbs *tænke*, *tenke* and *tänka* display the same possibility of non-experiential usage ('to intend', 'to consider', 'to reflect'). This is not the case for *synes, tro* or *tycka,* which are experiential verbs.

– Second, these three languages are all verb-second languages in declarative contexts. In other words, we can find the subject before example (8) or after example (9) the verb (here, *har* 'have').

(8) *Du har ca 14 dage til at bekræfte din bruger-konto.* (Danish)
you have approximately 14 days until to confirm you user-account
'You have approximately 14 days to verify your user account.'

(9) *Nu har du muligheden for et job på Danmarks hyggeligste arbejdsplads.* (Danish)
Nu har du mulighed-en for et job ...
Now you have possibility-DEF for a job
'Now you have the opportunity for a job at Denmark's cosiest workplace.'

The VS (verb-subject) order is possible in all non-subordinate contexts and mandatory in questions example (10).

(10) *Har du set denne person?* (Danish)
Have you seen this person
'Have you seen this person?'

The SV order, on the contrary, is possible only in subordinate clauses example (11) and in assertions.

(11) *Varen skal IKKE sendes retur til os før* du har modtaget *en mail retur fra os.* (Danish)
... *før du har modtaget en mail retur fra os* (Danish)
before you have received an mail back from us
'The item should NOT be sent back to us until you have received a mail answer from us.'

Since I want to examine possible biases towards assertion or interrogation, SV and VS tokens must be treated separately.

– Regarding the syntactic relationship between the mental verb and the clausal argument, we must first distinguish between syntactic disintegration (parentheticality proper) and syntactic integration with classical embedding, exemplified by the presence of a complementizer. Then, the distinctive "weakly embedding" structure with complementizer deletion must be treated separately as a potential bridge towards parentheticization. This means that we have to consider three great syntactic types: proper parentheticals (disintegrated), semi-parentheticals (lacking a complementizer) and matrix verbs overtly subordinating their propositional complement.
– Once this is done, we have to consider whether the epistemic verb stands before the content clause, after it or within the clause itself.
– Finally, the tokens must be annotated for the assertive vs interrogative value of the construction centred on the epistemic verb.

For each verb, four strings were isolated, distinguishing first and second person singular use and subject-verb and verb-subject order. For each of the four strings, 200 tokens were retrieved. These four types will now be designated as follows: S1-V, S2-V, V-S1, V-S2.

8.4 Results

8.4.1 Absolute frequencies in the corpora

First, let us turn to absolute frequencies. The following series of charts presents the ratio of verb-subject and subject-verb strings in the corpora for all verbs of the sample and for the verbal form *har* 'have', which I use as control item. *Har* is external to the field of epistemicity and can be used as an auxiliary for the perfect tense of most verbs in Scandinavian languages, so that it approximates what an unmarked distribution would be. The figures read as follows: if the ratio is above 1, the VS order is more frequent than the SV order. The lower the ratio is, the more dominant the SV order.

Already the first results are strikingly convergent in all three languages (Danish in Table 1, Norwegian Bokmål in Table 2, Swedish in Table 3): first, the second person (S2) always shows a slightly higher share of VS constructions than the first person (S1) does. The highest gap is observed on belief verbs (*tror* and *synes* / *syns* / *tycker*). For these verbs, the VS order even overrates the SV order in

Table 1: VS/SV ratio in Danish.

•	S1	S2	Total number
har	0.29	0.54	6,326,749
ved	0.36	0.62	756,465
tænker	0.55	0.82	205,869
tror	0.53	**2.31**	**864,482**
synes	0.57	**1.09**	**981,858**

Table 2: VS/SV ratio in Bokmål.

•	S1	S2	Total number
har	0.35	0.49	3,094,957
vet	0.33	0.89	385,149
tenker	0.77	0.69	206,968
tror	0.71	**2.22**	**839,011**
syn(e)s	0.77	**1.97**	**606,506**

Table 3: VS/SV ratio in Swedish.

•	S1	S2	Total number
har	0.50	0.68	5,141,630
vet	0.33	0.76	983,660
tänker	0.54	0.58	431,439
tror	0.60	**2.31**	**1,784,551**
tycker	0.71	**1.40**	**1,690.973**

the S2 row. At the same time, they are consistently the most frequent two verbs of the epistemic domain in the sample, ranging above KNOW (*ved*, *vet*) and the cognate of *think*, which is always the least frequent of the four. This is a strong sign that the Scandinavian cognates of *think* differ strongly from their West Germanic counterpart: In Scandinavian languages, *tænke-tenke-tänka* are not the standard opinion verbs. On the contrary, they seem to behave like marginal items in the set of epistemic predicates. This specific behavior of *think*-cognates will be observable in the following queries, too, suggesting that other lexical features compete with deictic mechanisms and can inhibit them.

In the case of *tror*, all three languages exhibit exhibit a ratio above 2, meaning that VS is more than twice as frequent as SV when the verb stands in the second person. By contrast, the cognates of *think* follow a different cline. In Norwegian Bokmål (Table 2), we find one exception to the rise of the ratio in the second person:

the ratio between Norwegian *tenker du* and *du tenker* is below the ratio between *tenker jeg* and *jeg tenker*.

8.4.2 Sample analysis (1): propositional scope

Turning to the samples made of 200 tokens per construction per verb, let us first examine the share of propositional scope in the whole set. Here, too, the cognates of *think* behave idiosyncratically, with a much lower share than the other two verbs taken in each language.

Table 4: Number of tokens with propositional scope in each sample of 200 tokens : Danish.

	S1V	VS1	S2V	VS2
tror	164	134	111	108
synes	195	166	156	88
tænker	81	75	65	102

A first glance at the numbers for Danish (Table 4) suggests that lexical idiosyncrasies are stronger than person-linked or order-linked regularities. This is corroborated by the calculation of the p-value of this distribution through an analysis of variance: p=0.6084 if we take the four constructions as parameters of observation; but if the variation criterion is the lexical item, p=0.0307, meeting the classical threshold of 0.05. Thus, the hypothesis that variation is dependent on the lexical entry is plausible, whereas an analysis taking word order or person as the determining factors does not lead to acceptable results.

The results in Norwegian (Table 5) and Swedish (Table 6) are highly consistent. The p value for the person+order hypothesis is of 0.5719 in Norwegian and 0.2105 in Swedish, corroborating the irrelevance of these factors. The lexical determination hypothesis is much more convincing, with p=0.0359 in Norwegian and p=0.0277 in Swedish.

Table 5: Number of tokens with propositional scope in each sample of 200 tokens: Norwegian Bokmål.

	S1V	VS1	S2V	VS2
tror	187	133	140	116
synes	183	169	166	70
tenker	66	110	19	73

Table 6: Number of tokens with propositional scope in each sample of 200 tokens : Swedish.

	S1V	VS1	S2V	VS2
tror	184	172	125	136
tycker	156	167	122	113
tänker	105	126	46	71

In Norwegian and Danish, the verb with the highest share of propositional scope is *synes*. At the other end of the spectrum, *tänka, tenke taenke* display a wide range of non-experiential uses, especially intentional ones.

(12) *Sådan har mine forældre også gjort, og jeg tænker at gøre det samme.* (Danish)
'That's also what my ancestors did, and I intend to do the same.'

These verbs also have the lowest share of wide-scope uses in the samples in Swedish and Norwegian. Both in Danish and in Norwegian, the gap between *tenker / tænker* and the other verbs narrows in the V-2SG construction, i.e. *tenker du / tænker du*. This is a first sign of a specific path of functional and formal specialization for this verb in the second person. Others are discussed below.

8.4.3 Sample analysis (2): share of direct questions

In what follows, I examine the frequency of each lexico-grammatical type in direct questions. I leave aside dependent interrogative clauses, where some properties of questions can be found, and where Scandinavian languages take S(-Adv)-V order. In direct questions, however, the main verb has to be preceded by the *wh*-item. For that reason, we can expect the figures to diverge dramatically between SV and VS configurations.

The general distribution in Table 7 (with Danish data) is significant (p=0.0076). After the analysis, person and word order appear to be approximately equally significant factors in a language like Danish. Examining variation only along the criterion of person, we find that p=0.0412; word order alone gives a p value of 0.0491, slightly less compelling than person.

The results from Norwegian (Table 8) and Swedish (Table 9) confirm that person and word order are on a par: P=0.0341 for person as a factor of variation in Norwegian (see the raw figures in Table 8) and 0.0120 in Swedish (Table 9). Word order is at 0.0293 in Norwegian (this time slightly more compelling than person), 0.0121 in Swedish.

Table 7: Number of tokens with a directive illocutionary force[4] in each sample of 200 tokens : Danish.

	S1V	VS1	S2V	VS2
tror	0	0	8	93
synes	0	0	3	22
tænker	0	3	0	77

Table 8: Number of tokens with a directive illocutionary force in each sample of 200 tokens : Norwegian Bokmål.

	S1V	VS1	S2V	VS2
Tror	0	1	14	95
synes	0	0	8	59
tenker	0	6	2	33

Table 9: Number of tokens with a directive illocutionary force in each sample of 200 tokens : Swedish.

	S1V	VS1	S2V	VS2
tror	0	0	4	136
tycker	0	0	5	113
tänker	0	0	0	59

8.4.4 Sample analysis (3): Position within the clause

8.4.4.1 Pre-clausal placement

Regarding the position before the clausal argument, at least two questions have to be raised:

(i) Of all person-order patterns, should we expect one of them to be over- or under-represented in the construction before the clausal argument? At first sight, we would not expect order to be a major factor, because the pre-clausal position can correspond either to a subject-verb cluster or to an X – verb – subject cluster, where X can be an anaphoric, deictic or argumentative adverb, for instance.

4 For reasons of terminological convenience, I use *directive* as a cover term for all flavors of questions and requests. Note that this would not necessarily be the case if we were also considering non-epistemic directive speech acts.

(13) *Så synes jeg det er et dårligt køb* (Danish)
So think I it is a bad purchase
'So I think it's a bad purchase.'

(ii) Are "weakly embedding", parenthetical constructions without a complementizer sensitive to person ? Or do they depend on lexical content ?

The first results for Danish (Table 10) suggest that significant regularities are observable (p= 0.0289):

Table 10: Strongly vs weakly embedding pre-clausal epistemic constructions in each sample of 200 tokens : Danish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror + COMPL	72	37	49	35
tror – COMPL	89	55	40	45
synes + COMPL	57	41	26	36
synes – COMPL	130	49	78	25
tænker + COMPL	50	28	31	24
tænker – COMPL	26	16	27	24

Because *tænker* seems to behave oddly, we can recalculate P for *tror* and *synes* only, and then p=0.0164. However, if we take the same figures but test them for lexical entries, we have p=0.0986. This means that lexical factors are relevant to isolate *tænker* from the other two, but are not a convincing factor to analyze the distribution as a whole.

Taking the strongly and weakly embedding constructions of all three verbs, neither person nor word order is a statistically significant factor on its own.[5] Without *tænker*, however, the distribution by person becomes narrowly significant (p=0.0512), whereas the p-value of the distribution by word order is almost unchanged at 0.0904. In other words: *tænker* does not behave differently from the other two with respect to word order, but it does differ from them when it comes to the selection of the grammatical person of the subject.

In Norwegian (Table 11), the general distribution has a p-value of 0.3182, which is not improved if we leave *tenker* aside (0.3382).

Thanks to further statistical analysis, it was possible to isolate one grouping where the distribution shows a convincing p-value: if we separate V+2SG from the other two and compare the first three columns, which are almost always instances of declarative or subordinate contexts, with the fourth one, where interrogative

5 P=0.1336 if we group the columns by person, P=0.0992 if we group them by word order.

Table 11: Strongly vs weakly embedding pre-clausal epistemic constructions in each sample of 200 tokens : Norwegian Bokmål.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror + COMPL	31	23	27	15
tror – COMPL	152	47	70	64
synes + COMPL	22	15	12	7
synes – COMPL	143	72	99	39
tenker + COMPL	49	45	35	20
tenker – COMPL	13	19	1	11

clauses have a much higher share, we end up with p=0.0349. If we leave aside the question of whether these verbs demand a complementizer or not, p is improved to 0.0229. However, this raises the question of what we are actually measuring: these results suggest that the use of these verbs in pre-clausal position is sensitive to illocutionary types, and not to person per se.

The only robust takeaway from this observation seems to be on the lexical side: *tenker* diverges from the other two verbs inasmuch as it is preferably constructed with a complementizer. The other two verbs strongly prefer the construction without a complementizer in any context. By contrast, in Danish, *tænker* is the only verb which systematically prefers the construction with a complementizer, but this does not mean that the other two systematically prefer the complementizer-less construction. The constructional idiosyncrasy of Norwegian *tenker* is much sharper than it is for Danish *tænker*. Swedish, finally, tells a different story, because the complementizer-less construction is systematically less frequent than the strongly embedding one (Table 12). Note, however, that a gap between *tänker* and its rivals can still be observed: *tänker* is the only verb where the complementizer-free construction never scores above 5 tokens per type.

Table 12: Strongly vs weakly embedding pre-clausal epistemic constructions in each sample of 200 tokens : Swedish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror + COMPL	111	71	70	82
tror – COMPL	69	28	24	8
tycker + COMPL	93	50	49	63
tycker – COMPL	48	35	32	16
tänker + COMPL	102	97	40	40
tänker – COMPL	1	0	5	1

Just like in Norwegian, there is no other significant takeaway from these figures (p=0.2518) except if we oppose the fourth row to the other three, with the same limit as before:[6] we are isolating a construction with a cluster of heterogeneous properties and "see" that its distribution is significantly different from a group that is defined as everything else, which is hardly useful.

Thus, as a whole, the results are inconclusive: it was not possible to isolate a positive influence of person on the distribution of epistemic verbal construction in pre-clausal position in Norwegian and in Swedish. Danish, on the other hand, could display such a deictic-egophoric dimension. There is, however, one consistent takeaway from these figures, which is the systematically lower range of syntactic parentheticization of *think*-cognates, which demand or prefer syntactic embedding with a complementizer.

8.4.4.2 Utterance-final position

In a Germanic V2 language, utterance-final epistemic verbs tend to show the VS order. This is also what we find in the sample, where there is not a single example of SV order in Danish or in Norwegian, and merely one in Swedish.

That being said, the observations do not necessarily lead to significant results. In Danish (Table 13), the results are strongly diverging (p=0.2433), leading to the hypothesis that variation is bound to strictly lexical factors. Once more, *tænker du* appears to be the item disrupting the picture:

Table 13: Utterance-final epistemic constructions in each sample of 200 tokens : Danish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	0	23	0	2
synes	0	8	0	5
tænker	0	24	0	51

This *tænker du* is almost always modified by an adverb or a particle (*nu* 'now', *så* 'so', *måske* 'maybe', *sikkert* 'certainly'). Its value mostly has to do with the anticipation of the addressee's reactions. Unsurprisingly enough, the written medium

6 In that case, p= 0.0139, improved to 0.0039 if we disregard the question of the complementizer. By comparison, in Danish, the same grouping would give p=0.0014, deteriorated to 0.0194 if we leave aside the question of the complementizer.

biases the usage towards something less interactional and more argumentative: 'think' here doesn't necessarily refer to something that the addressee may believe, but much rather to what they are likely to have in mind.

(14) *Ahem, men hvordan kommer alt dette så helt præcist til at have indflydelse på min virksomheds annoncering/annonceringsstrategi på Facebook,* tænker du måske? *Præcis hvilke ændringer der kommer til at ske og hvilke konsekvenser disse har for dig som annoncør/virksomhed.*
'Er, but how exactly will all this affect my company's advertising/advertising strategy on Facebook, *you might think?* Exactly what changes will occur and what consequences these have for you as an advertiser/company.'

In Norwegian (Table 14) and Swedish (Table 15), the results are more consistent (p=0.0085 for Table 14, 0.0012 for Table 15): the *think*-cognate is slightly less frequent in the second person than in the first, and the gap widens for the other two verbs. This suggests that *tænker* is further advanced in the pragmaticalization process in Danish than its cognates in the other two languages.

Table 14: Utterance-final epistemic constructions in each sample of 200 tokens : Norwegian Bokmål.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	0	31	0	7
synes	0	26	0	5
tenker	0	43	0	36

Table 15: Utterance-final epistemic constructions in each sample of 200 tokens : Swedish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	1	22	0	6
tycker	0	37	0	13
tänker	0	24	0	23

The final position is well-attested cross-linguistically for V-S(P1) (*tror jeg*, *synes jeg* etc. type). In average, there are around 20 tokens of utterance-final placement for each sample of 200. That position is only episodically attested for the V-S2 type (5 to 10 tokens / sample) except for *tänka* and its cognates (more than 20 tokens / sample, and even 51 tokens / 200 for *taenker du* in Danish). This use of *tænker du / tenker*

du / tänker du is almost exclusively interrogative: we can suspect that we are faced with a pragmaticalized question tag.

8.4.4.3 Clause-internal position

These items are also well-attested as parentheticals within the clause. This time, *tänka* and its cognates seem much less prone to clause-internal positioning than the other verbs. One should also note the fact that among the landing sites for these verbal constructs, the slot between the first phrase of the declarative clause and the finite verb is very common:[7]

(15) *Teori og empiri vil jeg gerne tale om. Resten* synes jeg *bliver ret useriøst.* (Danish)

. . . *Resten synes jeg bliver ret useriøst*
rest-DEF think I turns quite bad
'I would like to talk about theory and empiricism. The rest, *I think*, is quite frivolous.'

In Danish (Table 16), the results are crystal clear: no statistically significant relationship can be found if we take person or internal word order as determining factors (p=0.4656). However, if we test the hypothesis of a lexically-driven variation, we get a p-value 0.0428 :

Table 16: Clause-internal epistemic constructions in each sample of 200 tokens : Danish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	3	14	22	25
synes	9	43	47	20
tænker	3	5	7	1

Similarly, in Norwegian (Table 17), taking constructional patterns as drivers of variation yields non-significant results (p=0.3154), whereas lexically-driven variation is a much more robust hypothesis (p=0.0518).

7 This pattern is not specific for Scandinavian languages, and has also attracted the attention of Axel-Tober, Coniglio Müller andPaul (2025) in German.

Table 17: Clause-internal epistemic constructions in each sample of 200 tokens : Norwegian Bokmål.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	4	32	43	30
synes	18	55	55	20
tenker	2	2	13	6

The same holds for Swedish data (Table 18), with p=0.4611 if we take constructional features as drivers, but p=0.0276 if we re-calculate the results assuming that variation is lexically-driven.

Table 18: Clause-internal epistemic constructions in each sample of 200 tokens : Swedish.

	1sg + V	V + 1sg	2sg + V	V + 2sg
tror	3	37	31	39
tycker	15	43	41	20
tänker	0	5	1	5

8.5 Conclusion

Let us first summarize the results of this study.

– There is indeed a massive flip in frequency from almost exclusively declarative contexts in the first person singular towards a large share of interrogative contexts in the second person. The grammatical person proves to be a statistically significant factor. This finding corroborates the idea that there is a continuum between grammatical egophoricity and more general deictic features associated with the functional domain of epistemicity.
– The type of complementizer-less embedding that is typical for epistemic verbs and speech act verbs seems to be lexically-driven: not all opinion verbs are equally prone to embed their clausal argument without a complementizer. It is also not the same across Scandinavian languages.
– In all three languages, the cognates of English *think* are rather non-typical epistemic predicates: they exhibit a range of agentive, non-experiential uses and can also convey agent control over the thinking process. Conversely, these verbs also display the smallest level of syntactic and person-related idiosyncra-

sies. This strongly suggests that the semantic profile of the subject argument is a key factor at play: the interrogative flip depends on the subject being framed as an experiencer.

However, these *think*-cognates are also the verbs with the most distinctive pragmaticalized use in the second person. *Tænker du* (Danish), *tenker du* (Norway), *tänker du* (Swedish) are not only the most distinctive frozen or semi-semi-frozen second-person constructions in the corpus data at hand : they are also the most frequent epistemic units in utterance-final position. The final margin of the utterance seems to indeed be a privileged locus for discourse markers with intersubjective functions[8].

Overall, the results of this pilot study corroborate the claim that while Scandinavian language do not display egophoricity as a morphosyntactically relevant grammatical category, the behavior of epistemic verbs shows the same properties that were also observed on grammatical egophorics in other languages, most notably the interrogative flip. That being said, the precise characteristics of first- vs. second-person verbal constructs appear to be strongly dependent on lexical features.

Furthermore, even in languages as closely related as Danish, Swedish and Norwegian and on the basis of written corpora, these features are subject to significant variation from one language to the other. The deictic bias associating declarative contexts with first-person authority on opinion verbs and interrogative contexts with second-person authority may offer a key to approach language-specific phenomena, but it is not evenly manifested. One of these unequally relevant language-specific parameters is the constructional degree of parentheticization. However, in line with the argument formulated by Axel-Tober and Müller (to appear), the complementizer-free construction of epistemic verbs preposed to their clausal argument does not seem to be a credible locus for the incipient pragmaticalization of these predicates into discourse markers. On the contrary, the use of *tenker du* in late position is the most credible candidate to the role of an emerging discourse marker in the sample under consideration here. Consequently, the present study is also an invitation to shift the focus away from syntactic parentheticality and to regard epistemic verbs primarily as stance-markers with a tendency to pragmaticalization; the opposition between first and second person appears to be a major factor weighing on the pragmaticalization paths followed

8 This means that research on mental verbs or so-called “parenthetical verbs” may need to focus less on “parentheticality” as a syntactic property and less on first-person forms, too: assuming that epistemic predicates are intrinsically deictic, we should focus more on second-person forms.

by these verbs, corroborating the claim that categorial egophoricity supervenes on functional biases that are clearly manifested even in languages lacking conventionalized evidential and/or egophoric grammatical marking.

List of abbreviations used in the glosses

DEF	Definite
PAST	Past
PREP	Preposition
PRES	Present

References

Abraham, Werner & Elisabeth Leiss. 2012. Introduction: Theory of mind elements across languages. Traces of Bühler's legacy in modern linguistics. In Werner Abraham & Elisabeth Leiss (eds.), *Modality and Theory of Mind Elements across Languages*, 1–36. Berlin: De Gruyter.

Aijmer, Karin. 2011. I think — an English modal particle. In Toril Swan & Olaf Westvik (eds.), *Modality in Germanic languages*, 1–48. Place: De Gruyter.

Anscombe, Gertrude Elizabeth. 1957. *Intention*. Cambridge, Ma.: Harvard UP.

Axel-Tober, Katrin, Marco Coniglio, Kalle Müller & Katharina Paul. 2025. Grammaticalization of sentence adverbs and modal particles revisited. In Holly Kennard, Emily Lindsay-Smith, Aditi Lahiri & Martin Maiden (eds.), *Historical Linguistics 2022: Selected papers from the 25th International Conference on Historical Linguistics, Oxford, 1–5 August 2022*, 232–248. Amsterdam: John Benjamins.

Axel-Tober, Katrin & Kalle Müller. To appear. Sentence adverbs based on parenthetical clauses – the Case of German. In Marco Coniglio, Kalle Müller & Markus Steinbach (eds.), *Adverbs and Particles at the Form-Meaning Interface*. Amsterdam: John Benjamins.

Beeching, Kate & Ulrich Detges (eds.). 2014. *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change*. Leiden: Brill.

Bergqvist, Henrik. 2021. Egophoricity and Perspective: A view from spoken Swedish. *Frontiers in Communication* 6. 627144.

Bergqvist, Henrik & Seppo Kittilä. 2019. Epistemic perspectives: Evidentiality. egophoricity, and engagement. In Henrik Bergqvist & Seppo Kittilä (eds.), *Evidentiality, egophoricity and engagement*, 1–21. Berlin: Language Science Press.

Bergqvist, Henrik & Karolina Grzech. 2023. The role of pragmatics in the definition of evidentiality. *STUF – Language Typology and Universals* 76/1. 1–30.

Blanche-Benveniste, Claire. 1989. Constructions verbales en incise et rection faible des verbes. *Recherches sur le français parlé* 9. 53–73.

Boye, Kasper & Peter Harder. 2021. Complement-taking predicates, parentheticals and grammaticalization, *Language Sciences* 88. 101416.

Dahl, Östen. 2000. Egophoricity in discourse and grammar. *Functions of Language* 7. 37–77.

Davidson, Donald. 1984. First-person authority. *Dialectica* 38/2–3. 101–111.
Dehé, Nicole & Yordanka Kavalova. 2007. Parentheticals: an introduction. In Nicole Dehé & Yordanka Kavalova (eds.), *Parentheticals*, 1–22. Amsterdam: John Benjamins.
Faller, Martina. 2024. The interrogative flip with illocutionary evidentials. *Folia Linguistica* 59. [ahead of print, https://doi.org/10.1515/flin-2023-2050]
Günthner, Susanne & Wolfgang Imo. 2003. Die Reanalyse von Matrixsätzen als Diskursmarker: *ich mein*-Konstruktionen im gesprochenen Deutsch. *Inlist* 37. 1–35.
Hargreaves, David. 2018. Am I blue? Privileged access constraints in Kathmandu Newar. In Simeon Floyd, Elisabeth Northcliffe & Lila San Roque (eds.), *Egophoricity*, 79–108. Amsterdam: John Benjamins.
Hassler, Gerda. 2011. Epistemic modality and evidentiality and their determination on a deictic basis. In Gabriele Diewald & Elena Smirnova (eds.), *Linguistic realization of evidentiality in European languages*, 223–248. Berlin: De Gruyter.
Heine, Bernd & Günter Kaltenböck. 2021. From clause to discourse marker: on the development of comment clauses. *Language Sciences* 87. 101400.
Jakobson, Roman. 1971 [1957] Shifters, verbal categories, and the Russian verb. In Roman Jakobson, *Selected writings volume II, Word and language*, 130–147. The Hague: Mouton.
Leiss, Elisabeth. 2012. Epistemicity, evidentiality and Theory of Mind. In Werner Abraham & Elisabeth Leiss (eds.), *Modality and Theory of Mind Elements across Languages*, 39–66. Berlin: De Gruyter.
Martineau, France. 1993. Rection forte et rection faible des verbes: l'ellipse de Que en français du Québec et de l'Ontario. *Francophonies d'Amérique* 3. 79–90.
Mitchell, Keith. 2009. Semantic ascent, deixis, intersubjectivity and modality. In Raphael Salkie, Pierre Busuttil & Johan van der Auwera (eds.), *Modality in English: Theory and Description*, 55–78. Berlin/ Boston : Mouton de Gruyter.
Modicom, Pierre-Yves. 2025. *Subject properties in Early Modern Germanic Languages: A contrastive corpus-based study.* Berlin: De Gruyter.
Russell, Bertrand. 1910. Knowledge by acquaintance and knowledge by description, *Proceedings of the Aristotelian Society* 11. 108–128.
San Roque, Lila, Simeon Floyd & Elizabeth Northcliffe. 2017. Evidentiality and interrogativity. *Lingua* 186–187. 120–143.
San Roque, Lila, Simeon Floyd & Elizabeth Northcliffe. 2018. Egophoricity: An introduction. In Simeon Floyd, Elizabeth Northcliffe & Lila San Roque (eds.), *Egophoricity*, 1–78. Amsterdam: John Benjamins.
Schiffrin, Deborah. 1987. Information and participation: Y'know and I mean. In Deborah Schiffrin, *Discourse Markers*, 267–311. Cambridge: Cambridge UP.
Urmson, James Opie. 1952. Parenthetical Verbs. *Mind* 61(244). 480–496.
van Olmen, Daniël & Jolanta Šinkūnienė. 2021. Pragmatic markers and peripheries: An overview, in Daniel van Olmen, Daniël & Jolanta Šinkūnienė (eds.), *Pragmatic Markers and Peripheries*, 1–16. Amsterdam: John Benjamins.