

Grammatical person as a clue to interpreting French how-questions

Jan Fliessbach¹, Lucia M. Tovenà², Damien Fleury²,

¹University of Potsdam, ²Université Paris Cité

Correspondence: jan.fliessbach@uni-potsdam.de

Abstract

The French wh-phrase *comment* (how/why) contributes only in a limited way to constraining the interpretation space of the questions in which it appears. It may refer either to a manner, means, or method, understood as properties of events, or to a reason, understood as a semantic proposition. This paper presents a corpus study investigating how the readings of *comment* questions in dialogue correlate with the grammatical person and number of the subject. We find that manner readings typically involve third-person subjects, method readings are associated with first-person subjects, and reason readings occur predominantly with second-person singular subjects. We interpret these correlations as reflecting key properties of the dialogical context, supporting a view of person as a discourse-sensitive indicator of epistemic roles. These associations between person and *comment* readings are compatible with broader cross-linguistic patterns involving egophoricity and evidential access, despite the fact that French does not have dedicated morphological markers for these categories.

1 Introduction

Speakers use questions in dialogue to indicate their goal of requesting specific information. For a question to be used felicitously, the speaker must lack the relevant information and believe that the addressee might possess it (cf. Searle, 1969, and subsequent work). Beyond recognizing questions as such, the correct classification of the interpretation space of questions (i.e. identifying which piece of information is requested) plays a pivotal role in verbal interaction. Wh-phrases contribute to this process through their intrinsic characterisation, e.g., [+human] for Fr. *qui* and En. *who*, [+time] for Fr. *quand* and En. *when*, etc. However, wh-phrases such as Fr. *comment* and En. *how* vary between manner, means and method, qualifying as properties of events (Sæbø, 2016), and reason, qualifying

as semantic propositions (Tovenà, 2023). In example (1), *comment* asks for manner [good, bad, . . .], in (2) for a method [by getting onto hands and knees, by showering, . . .] or means [with grab bars, with a bar lift, . . .], and for a reason in (3) [coincidences, lack of trust, . . .] (glosses in Appendix A).

- (1) a. MITCH: **Comment va l'enfant ?**
(How is the child?)
b. JULIA: Il va bien.
(He's fine.) (Julia, 2008)
- (2) Context: Marthe asks Lulu to help her get out of the bathtub.
a. LULU: **Comment vous faites quand vous êtes toute seule ?**
(How do you manage when you're all alone?)
b. MARTHE: Je prends pas de bain, c'est tout. Un peu de nerf ma grande.
(I don't take baths, that's all. Get a grip, girl.) (Lulu femme nue, 2013)
- (3) Context: Denis Robert suspects Imad Lahoud of passing on information.
a. IMAD LAHOUD: **Comment tu peux imaginer que je puisse être assez fou pour avoir fait ça ?!**
(How can you imagine I would be crazy enough to have done that?!)
b. DENIS: Comment t'expliques ces coïncidences, alors ?
(So how do you explain these coincidences?) (L'enquête, 2014)

The focus of this study lies in charting which grammatical person and number marked on the subject is preferably associated with which reading of *comment*¹ questions. Grammatical person is

¹On the interpretation of this wh-phrase, see Olivier (1985); Moline (2009); Fleury and Tovenà (2018); Van de Velde (2009), among others.

marked overtly or easily reckoned, while deducing the interpretation of the *comment*-interrogative is less straightforward. Thus, the possibility of making reliable inferences from the person and number of the subject to the meaning of the question offers a plausible strategy for interlocutors navigating cognitive and temporal constraints in interaction, and may also inform the design of dialogue systems.

2 Background

The grammatical category of person covers the expression of the distinction between i) the speaker of an utterance, referred to as the first person, ii) the addressee, or second person, and iii) any referent who is neither the speaker nor the addressee, known as the third person (Siewierska, 2004, 1). All three persons have distinct singular and plural number in French, with the second-person plural also used for polite singular reference (*vouvoiement*).

When utterances are considered within the context of conversation, the classification must be expanded to include the notion of discourse role. The roles of speaker and addressee correspond to the participants in a conversation and are typically referenced by first- and second-person singular forms—though notable exceptions include quoted speech, inner dialogue, and generic uses of the second person. In root declarative sentences, the speaker is commonly the source of semantically determined information. In information-seeking questions, however, the perspective shifts in what is often called interrogative flip. The expected source of knowledge (termed the *assertor* by Creissels 2008) is typically the addressee. Intuitively, a question like *Who did you see?* requires far less contextual support than *Who did I see?*, the latter implying that the speaker either lacks access to their own perceptual experience or is testing whether the hearer knows the answer (a so-called quiz-question).

In addition to discourse and speech act related roles of speaker and addressee, the event and its participants have an impact on structuring discourse. Person and number agreement marking on the predicate reflects grammatical information and thematic roles, in French and other languages.

Note that event related information is exploited beyond agreement marking. The experiencer holds a special epistemic status, and languages have various means of foregrounding the holder of first-hand knowledge or marking specific viewpoints. This broader perspective encompasses *egophoric* sys-

tems, which encode a distinction between the *assertor* and all others, and are found in languages that typically do not mark the grammatical agreement of the person on the verb (San Roque et al., 2018, 49). *Egophoric* forms encode privileged access to the information conveyed in an utterance. Marking is typically done on the predicate and tends to be restricted to assertors with an active involvement (Creissels, 2008). Markers interface with the “traditional” roles of speaker and addressee, usually occurring in restricted distribution according to speech act role and sentence type. This person sensitive distribution reflects principles of epistemic authority to judge the truth of the proposition, access, and potential self-ascription (San Roque et al., 2018, 2). In contrast, *allophoric* forms signal an external perspective lacking that authority. The prototypical distribution of egophoric markers is illustrated in Table 1, adapted from Widmer and Zúñiga (2017, 420).

Table 1: Typical distribution of egophoricity markers

person	assertion	question
1	ego	allo
2	allo	ego
3	allo	allo

The marking of the holder of epistemic knowledge, mainly on the predicate, is found in another phenomenon, called *evidentiality*. The fundamental difference between the two is that evidentiality encodes the source of information, while egophoricity encodes the epistemic status or quality of the speaker’s knowledge (Floyd et al., 2018; Bergqvist and Kittilä, 2019). Although certain information sources may pragmatically extend to certain epistemic statuses across languages, e.g. visual perception usually yielding more reliable knowledge than hearsay, the distinction remains central to our aim of understanding how person interacts with the interpretation of *comment* questions because the different readings (manner, method, means, reason) have semantic components that require different sources (sensory vs. mental) and involve different configurations of epistemic authorities among speaker and addressee (ego- vs. allophoric).

HOW questions in general, with French *comment* (how/why) among them, allow for a wide range of possible answers. Following Sæbø (2016), we assume that *comment* denotes properties of events in a neo-Davidsonian sense. More precisely, the predicate of a HOW question denotes a set

of events, and when combined with (the trace of) *comment* in its manner reading, the result is the intersection of this set with the extension of *comment*. In this case, manner functions as an adjunct, and its presence does not alter the logical representation of its sister node. By contrast, in method readings, *comment* introduces an argument rather than an adjunct. This argument saturates a function, thereby altering the predicate’s semantic type: it now denotes a function from properties of events. Combining this with (the trace of) *comment* yields the application of that function to the intension of the trace. Finally, a third case is that of HOW with reason reading. Here, the predicate forms part of a saturated proposition rather than a propositional function, since there is no dependency between the wh-phrase *comment* and a trace within the clausal core. The wh-phrase denotes a function that, when applied to the prejacent (i.e. the proposition expressed by the clausal core), yields a set of propositions logically related to it in a way that supports a discourse-level explanation (Tovena, 2023).

The reason reading of *comment* is typically associated with surprise or epistemic misalignment (Fleury and Tovena, 2018). This is related to the notion of *mirativity*, a category that marks information as not yet integrated into the speaker’s knowledge state or as involving a form of psychological distancing (cf. DeLancey, 1997; Lazard, 1999; Bickel, 2008). Three entities are particularly relevant in reason *comment* questions. First, the speaker, who expresses surprise and whose expectations are contradicted by the prejacent. This is the facet of the reading that brings it close to mirativity. The ‘scope proposition’ discussed in formal accounts of mirativity and evidentiality corresponds to the prejacent in this context. Second, the addressee typically serves as the epistemic authority, as in canonical questions (Farkas, 2022). Third, the subject of the clause expressing the prejacent may be the speaker, the addressee, or a third party. Since it is never the gap or trace related to *comment*, its referent is always accessible and may be the source of epistemic knowledge about the described event. When the subject is marked for the first person and the verb is in the active voice, the configuration is likely to correspond to the case of direct access identified by Garrett (2001, 105), and to what Tournadre (2008) and Widmer and Zúñiga (2017) describe as involving an endopathic experiencer “in a state of affairs that involves a mental state or process that is only directly accessible to

the experiencer herself or himself” (e.g., *hungry*, *exhausted*) (Widmer and Zúñiga, 2017, 433).

Importantly, the correspondence between the speaker and the first-person singular subject (except in cases of quoted speech and inner dialogue), and the hearer and the second-person singular subject (except in generic uses), sets these two apart from the third person and from the plurals, particularly with regard to which sources of information are available to them. Here, again, we can draw on a comparison with languages that grammatically mark information source. Well-known models of evidentiality, e.g., Aikhenvald (2004) or Hengeveld and Hattner (2015), allow us to distinguish up to seven types of access to information: participatory, visual, sensory, inferential, presumptive, hearsay, quotative. We adopt a simplified ternary distinction between sensory (participatory, visual, sensory), mental (inferential, presumptive), and verbal (hearsay, quotative) information source to account for the differences between the readings of *comment*. As detailed in Section 3.3, method, means and reason questions target abstract relations between facets of events and goals or expectations, which are not directly accessible merely via the senses. Instead, such questions require mental processes: linking actions to goals (method, means) or integrating new information with expectations and logical entailments (reason). As a result, sensory input constitutes a lower-quality information source for addressing these more complex questions.

While French does not encode evidentiality or egophoricity morphologically, the upshot of this background section is that person and number marking reflects event participation but also access to mental representations of goals and expectations, thereby restricting the possible sources of information on which an answer can be based. We may therefore plausibly hypothesise that grammatical person and number, as encoded in French morphosyntax, correlate with specific readings of *comment* questions. Section 3 presents a corpus study designed to test this hypothesis by examining how different person-number combinations pattern with the distinct interpretive categories—manner, method, means, and reason—discussed in the preceding sections. Our analysis seeks to determine whether the absence of dedicated morphological markers is compensated for by systematic patterns in the use of person and number in *comment* questions that reflect underlying epistemic or discourse-related structures active across languages.

3 Corpus study

This section reports a corpus study on the readings of *comment* questions extracted from a corpus of 99 film scripts sourced from *Lecteurs Anonymes* (Fliessbach et al., 2024). Below, we detail our extraction and annotation methods, followed by the results of our statistical analysis.

3.1 Data and methodology

Given the unavailability of large, genre-homogeneous corpora of conceptually oral (Koch and Oesterreicher, 1985), informal French dialogues (see Fliessbach and Rockstroh 2024), we used film scripts because they have been shown to successfully approximate dialogue (Levshina, 2017; Bednarek, 2018). We relied on AntConc (Anthony, 2018) to extract occurrences of *comment*, including their immediate discourse context, from our corpus, discarding cases with no person marking (infinitive or verb-less questions). 786 occurrences could be annotated for both reading type and person. The authors were supported by a linguistically trained native speaker of French, who also added a blind control to 454 of the authors' reading annotations.² In cases of intuition mismatch, another annotation by one of the authors (blind to the others) was added to decide on the final reading by majority. Agreement between the control and the corresponding subset of final annotations was 87.2%; Cohen's $\kappa = 0.82$ (calculated with the *irr* package in R; Gamer et al. 2019; R Core Team 2025). Reason readings showed the lowest agreement (77%) (cf. Table 5 in Appendix B).

Initially, we identified six readings: MANNER, REASON, METHOD, MEANS, CLARIFICATION REQUEST (CR), and OTHER. However, due to the metalinguistic nature of CRs (Purver, 2004), and the low number of OTHER uses, we discarded both categories ($n=27$). All combinations of person (1,2,3) and number (singular and plural) in the French inflectional system are attested in the remaining sample ($n=759$), as well as impersonal constructions, in which pronouns such as *on* and *ça* do not refer to a specific entity. We found 87 polite plural forms for singular addressees and no third-person address uses. We included human and non-human subject referents, the latter being of particular importance among the third-person subjects.

²We thank Yoan Linon for the support and our reviewers for suggesting additional agreement measures.

3.2 Results

Tables 2 and 3 display the distribution of *comment* readings across person and number categories. The most frequent reading in the dataset is MANNER, followed by MEANS, METHOD, and REASON. In terms of grammatical person, second-person subjects occur most frequently ($n=273$), followed by third- ($n=243$) and first-person ($n=121$), as well as impersonal constructions. Plural forms (excluding politeness *vous*) are significantly less frequent than singular forms. The cross-tabulation of person and number reveals an overall even distribution of singular and plural across person categories, with a slight over-representation of third-person plural. Table 4 differentiates *tutoiement* from the polite plural *vouvoiement*.

Table 2: Person & number by reading of *comment*

	man	mea	meth	rea	Sum
1s	19	18	40	19	96
2s	78	61	50	56	245
3s	113	48	26	19	206
1p	5	6	13	1	25
2p	11	10	6	1	28
3p	13	10	12	2	37
imps	57	23	27	15	122
Sum	296	176	174	113	759

Table 3: Number categories by reading of *comment*

	man	mea	meth	rea	Sum
s	210	127	116	94	547
p	29	26	31	4	90
imps	57	23	27	15	122
Sum	296	176	174	113	759

Table 4: Address by reading of *comment*

	man	mea	meth	rea	Sum
vousPol	29	23	16	19	87
tu	49	38	34	37	158
other	218	115	124	57	514
Sum	296	176	174	113	759

Figure 1 shows the Pearson residuals from a χ^2 test of the distributions in Table 2. Figures 2 and 3 in Appendix C provide such information for Tables 3 and 4. Bar width reflects frequency, while color encodes the direction and significance of deviation from expected values: blue indicates significantly positive association, red indicates significantly negative association. All plots were generated using the *vcd* package (Meyer et al., 2006; Zeileis et al., 2007).

Each singular person category is positively correlated with one reading of *comment*: third-person singular subjects are associated with manner readings, first-person subjects with method readings (singular and plural), and second-person singular subjects with reason readings. Among the aggregated number categories (Appendix C), plural forms are negatively associated with reason readings and positively with method readings. Reason readings are positively associated with informal *tutoiement*. Impersonal subjects do not show a significant association with one of the readings.

More globally, the tendencies for the first-person singular complement those for the third-person singular. In particular, the first-person singular is negatively associated with manner and positively associated with method readings, two tendencies which are inverted for the third-person singular. The second-person singular, on the other hand, aligns more closely with the first-person singular, but deviates in its stronger association with reason readings.

3.3 Discussion

In the following subsections, we will discuss the associations between the readings and the person and number categories detailed in Section 3.2.

3.3.1 Manner

The most frequent interpretation of *comment* in our data is the manner reading, which shows a significant positive association with third-person singular subjects and a negative association with the first-person singular. This pattern can be explained in terms of epistemic access and aligns with the tendency of third-person-subject questions to receive allophoric marking. In questions with third-person subjects, the individual referred to is typically distinct from both speaker and addressee. As a result, the source of knowledge about the manner in which an event occurred is external to the discourse participants. Since manner information (e.g., how an action was performed) is often accessible through external observation, third-person subjects are well-suited for such readings.

Third-person subjects in our data frequently involve non-human referents, which are less compatible with reason or method readings. The former presupposes expectations; the latter, intentional goal-directedness (cf. Section 3.3). Non-human subjects thus naturally align with manner readings, where observable behaviour is more salient than

internal rationale or purposeful planning.

As for the negative association between first-person singular subjects and manner readings, this can be attributed to the implausibility of the speaker inquiring about properties of an event that they experience. For the second-person singular, our data contain situations in which speaker and hearer are co-present and both experience the properties of present and ongoing events, which reduces the likelihood of an inquiry into the event properties. However, there are cases in which co-presence of the interlocutors does not ensure mutual access to the relevant event properties. In our results, many of the exceptions to the negative association between first person and manner involve questions in which the speaker's attention is directed toward how they are perceived externally. Example (4) involves a question about the speaker's appearance or an externally visible characteristic. In such cases, the addressee can plausibly be better positioned to access the relevant information, making a manner reading with a first-person subject both possible and natural.

- (4) Context: A dancer asks their trainer after a performance:
J'étais comment ? (How was I?)

In sum, the fact that manner questions show a significant positive association with third-person subjects in our dataset can be accounted for by the observation that external, sensory access is necessary (and often sufficient) to provide information on how an action was performed.

3.3.2 Method and means

Method readings of *comment* show a significant positive association with first-person subjects, and a negative association with third-person subjects. Means readings do not show a significant effect. Lexically, method questions often involve verbs such as *faire (pour)* 'do/go about sth', *s'y prendre* 'go about sth', *atteindre* 'reach/achieve' where the method has argument status, and the subject is promoted to the role of a sentient agent volitionally aiming to cause an event or a change of state (Dowty, 1991, 572). Example (5) is a revealing case: it expresses the speaker's perplexity upon finding themselves in a situation where their usual methods are unavailable. The addressee is cast as an authority figure who is responsible for blocking the speaker's methods in the current context and

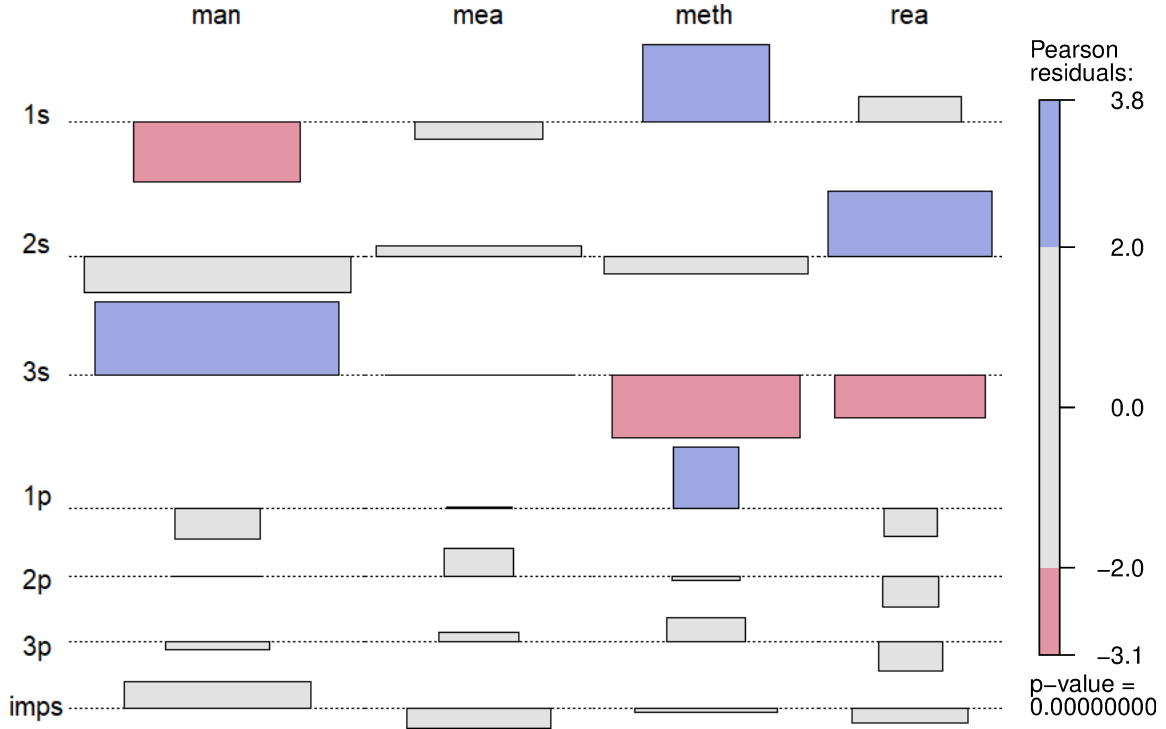


Figure 1: Association plot based on a χ^2 test on Table 2 (person/number categories by readings).

expected to suggest alternative solutions that would allow the subject to actualize a future event.

- (5) **Comment je fais si y’a rien ici ?** On l’opère au couteau sans anesthésie ? Je vais le tuer votre mec.
 (How do I manage if there’s nothing here? Do we operate him with a knife and without anesthesia? I’m going to kill him, your guy.)
 (Nos résistances, 2009)

At first sight, the nature of the relevant information on a method to perform an action appears objective and independent of the performer. Such information is not *per se* privileged knowledge held by a given discourse participant. However, it is worth noting that method questions typically license complete answers regarding abstract predicates (Sæbø, 2016, 6-7). As discussed in Section 2, such abstract information requires mental rather than sensory access. The information needed to determine a suitable answer may involve the speaker’s intentions and their internal reasoning. These questions thus presuppose that the referent of the grammatical subject had some internal motivation or strategy, making first-person subjects particularly compatible. Strictly speaking, the addressee is not required to share the speaker’s reasoning nor to be aware of it. As usual, the question arises from the speaker’s need to obtain some missing information,

paired with the assumption that the addressee can provide it in the given context.

What is less usual is that the information to be provided is about an action to be performed by the speaker. A first-person question would typically be marked as allophoric in languages with grammatical egophoricity. However, if the speaker’s goals are in question, introspection by the speaker might be the privileged source of information. By contrast, third-person subjects will usually not provide information to which they have introspective access, since they are not addressed by the speaker (in the second person). Furthermore, method and means readings do not target reasons or justifications in the epistemic sense (as reason readings do), but rather the operational path by which an outcome is to be achieved. This kind of procedural inquiry fits naturally with the speaker soliciting one *possible* option, not all the options that the addressee can think of.

In essence, the significant positive association of method readings with first-person subjects found in our dataset does not violate the principles that guide egophoric marking in other languages. A speaker who is also the referent of the grammatical subject has privileged access to the internal motivation or strategy in planning the action and can best assess what information is missing to actualise it.

Finally, the positive association between first-person plural subjects and method readings is due to questions regarding future communal endeavors, which invite group members to deliberate on strategies in a manner reminiscent of hortative modality.

3.3.3 Reason

The reason reading of *comment* is typically associated with epistemic misalignment or surprise, where the speaker seeks an explanation to reconcile the prejacent (i.e. the proposition *tu es aussi naïf* in example (6), *tu sais qu'elle boit* in (7b)) with their expectations or epistemic state (Fleury and Toven, 2018). As noted in Section 2, *comment*-reason-questions do not contain a low syntactic gap within the clausal core. Instead, the values of the *wh*-phrase are full propositions, which is why these questions are understood as querying explanations, rather than event-internal properties.

- (6) **Comment peux-tu être aussi naïf ?** Tu ne vois donc pas que les hommes qui te prennent le font pour leur propre plaisir ?
(How can you be so naïve? Can't you see that the men who take you are doing it for their own pleasure?)
(Les rencontres d'après minuit, 2013)
- (7) a. GRAND PAPY:
N'importe quoi... Elle dit n'importe quoi ! Ça quand on boit, on déraile.
(Nonsense... She's talking nonsense! When you drink, you go off the rails.)
- b. PETIT SPIROU:
Comment tu sais qu'elle boit ?
(How do you know that she drinks?)
(Le petit Spirou, 2016)

As discussed in Section 2, such readings presuppose mental access to a proposition rather than sensory access to event-properties. Differently from declarative miratives, the interrogative flip in reason questions places the speaker in the role of information-seeker, and the addressee is treated as the epistemic authority. When the subject is second-person singular, it is usually the addressee, who is human, individuated, and discourse-proximate, and can thus plausibly access the reasoning behind the proposition being questioned. This might account for the observed positive association between second-person singular subjects and reason readings, and the significant negative association with plural subjects, who typically represent distributed

knowledge and epistemic authority, and therefore lack individualized epistemic grounding.

While third-person singular subjects are not excluded, they are less likely in reason readings. Instead, the third-party referent typically serves as the target of the speaker's incredulity or reproach, rather than its resolution. It is still the addressee who is positioned to explain or justify the described state of affairs. Example (8) illustrates this: the third-person plural subject (*ils*) is not expected to provide an answer; the speaker appeals to the co-present addressee for interpretive alignment.

- (8) PAUL: Ils l'ont exclu du truc !! Tu comprends ? Mon fils. Ça vient de moi cette idée !!! **Comment ils ont pu lui faire ça ?**
Il sera jamais accepté dans cette école...
(They kicked him out of the thing!! You get it? My son. That idea came from me!!! How could they do that to him? He'll never get accepted to that school...)
(La lutte des classes, 2017)

First-person singular subjects are rarer but not absent. These cases raise a conceptual tension: the first person marks the speaker as event participant, a role usually associated with direct epistemic access, which would make such questions infelicitous. However, when the speaker's authority is undermined, e.g., due to perceived incompetence, reason readings remain available. These questions often involve modals of ability, as in (9), suggesting surprise at one's own actions and inviting a form of self-directed justification.

- (9) **Comment j'ai pu ne rien voir... comment j'ai pu me tromper comme ça ?**
(How could I not have seen... how could I be so wrong?)
(En liberté, 2017)

These utterances occupy a hybrid space between internal reflection and public address, consistent with cross-linguistic accounts of mirativity, where first-person subjects appear in contexts of epistemic dissociation (e.g., drunkenness or sleep, DeLancey 1997, 35). Indeed, HOW/WHY questions of reason carry nuances of reproach cross-linguistically, although the degree of negative stance by the speaker towards the state of affairs varies and can range from mild puzzlement to disapproval or denial (Toven, 2023). These nuances can also be related to the association between *tutoiement* and reason readings illustrated in Figure 3 in Appendix C: using an interrogative as a reproach might constitute

a face-threatening act, thereby undermining the politeness function typically associated with *vouvoisement*.

Another form of speakers distancing themselves from the state of affairs, and associating the audience in the same stance, is observed in a rhetorical reason question such as (10).

- (10) **Comment je pourrais ne pas être jaloux?**
(How could I not be jealous?)
(Le Mystère Henri Pick, 2018)

These first-person questions often involve modals of ability, highlighting the speaker's perceived failure or lack of competence. Finally, a form of rejection is also found in cases where the first-person question is used to challenge an attributed statement, as in reported speech (11).

- (11) **C'est quoi ça? Comment je paie la commande ?**
(What is that? What do you mean I pay for the order?)

In sum, the distribution of reason readings supports the broader hypothesis, put forward in Section 2, that these questions require mental access, which favors subjects with individuated mental states, most notably the second-person singular. First-person uses become felicitous when the speaker's epistemic stance is compromised or performatively framed. Third-person and plural subjects, lacking such individualized mental access, are generally dispreferred. This supports the claim that *comment* reason readings are shaped by epistemic role asymmetries and the mental accessibility of the justification being sought.

4 Conclusions

This study has examined how the French wh-phrase *comment* (how/why) interacts with grammatical person and number in shaping the interpretation space of questions. While *comment* can vary between manner, method, means, and reason readings, our corpus data reveal systematic correlations between these interpretations and the person-number features of the subject. Third-person subjects are associated with manner readings, indicating an allophoric perspective in which knowledge about the event is derived from external observation and attributed to the addressee (as the assertor), not to the third-person subject referent. In contrast, method readings are more frequent with

first-person subjects, consistent with the idea that such questions inquire about information accessible via mental processes (introspection, inference). Finally, reason readings are associated with second-person singular subjects, underscoring the role of the addressee as the epistemic authority in resolving violations of the speaker's expectations.

These findings support a view of person not merely as a deictic category, but as a discourse-sensitive indicator of epistemic roles. These associations between person and *comment* readings are compatible with broader cross-linguistic patterns involving egophoricity, mirativity, and evidential access, even though French lacks dedicated morphological markers for these categories. In this sense, the use of the grammatical person system in French reflects distinctions of knowledge access and assertoric authority that are grammatically encoded in other languages. Future research might examine whether similar patterns appear in further languages lacking dedicated epistemic or egophoric morphology. The interface between question type, subject properties (e.g. animacy or definiteness), active involvement in a state of affairs, and perspective-taking mechanisms also remains a promising domain for cross-linguistic inquiry. Finally, future research could investigate whether these person–reading correlations reflect more fundamental, cross-linguistic speaker strategies for managing epistemic asymmetries (Heritage, 2012). Speakers identify knowledge gaps and choose addressees based on assumptions about who is best positioned to provide the missing information. Investigating how such epistemic considerations shape the mapping between question type, subject person, and expected answer content could illuminate broader inquisitive strategies underpinning the observed patterns.

Acknowledgements

This work has been partially funded by Campus France and the German Academic Exchange Service (DAAD), Grant ID 57701768. We again thank our reviewers and Yoan Linon.

References

- Alexandra Y. Aikhenvald. 2004. *Evidentiality*. [Oxford linguistics]. Oxford Univ. Press, Oxford.
- Laurence Anthony. 2018. [AntConc 3.5.7 \[Software\]](#).
- Monika Bednarek. 2018. *Language and television se-*

- ries: *A linguistic approach to TV dialogue*. Cambridge Univ. Press, Cambridge.
- Henrik Bergqvist and Seppo Kittilä, editors. 2019. *Evidentiality, egophoricity and engagement*. Language Science Press, Berlin.
- Balthasar Bickel. 2008. Verb agreement and epistemic marking: a typological journey from the Himalayas to the Caucasus. In Brigitte Huber, Marianne Volkart, and Paul. Widmer, editors, *Chomolangma, Demawend und Kasbek*, pages 1–14. Int. Inst. for Tibetan and Buddhist Studies, Halle.
- Denis Creissels. 2008. [Remarks on so-called “conjunct/disjunct” systems](#). Paper presented at the 3rd Syntax of the World’s Languages Conference, Berlin, 25–28 September.
- Scott DeLancey. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology*, 1:33–52.
- David Dowty. 1991. Thematic proto-roles and argument selection. *Language*, 67(3):547–619.
- Donka F. Farkas. 2022. Non-intrusive questions as a special type of non-canonical questions. *Journal of Semantics*, 39(2):295–337.
- Damien Fleury and Lucia M. Tovenà. 2018. [Reason questions with *comment* are expressions of an attributional search](#). In *Proceedings of SEMDIAL 22 (AixDial)*, pages 112–121.
- Jan Fliessbach, Lisa Brunetti, and Hiyon Yoo. 2024. [On the overlapping discourse functions of Spanish ‘cómo que’ and French ‘comment ça’ interrogatives](#). In Külli Habicht, Tiit Hennoste, Helle Metslang, and Renate Pajusalu, editors, *Subjectivity and Intersubjectivity in Language*, Open Linguistics. de Gruyter, Berlin.
- Jan Fliessbach and Johanna Rockstroh. 2024. [Verb influence on French wh-placement: a parallel corpus study](#). *Corpus Linguistics and Linguistic Theory*, (1).
- Simeon Floyd, Elisabeth Norcliffe, and Lila San Roque. 2018. *Egophoricity*. John Benjamins, Amsterdam.
- Matthias Gamer, Jim Lemon, Ian Fellows, and Puspendra Singh. 2019. [irr: Various coefficients of interrater reliability and agreement](#).
- Edward John Garrett. 2001. *Evidentiality and assertion in Tibetan*. Ph.D. thesis, Univ. of California, LA.
- Kees Hengeveld and Marize Mattos Dall’Aglio Hattnher. 2015. [Four types of evidentiality in the native languages of Brazil](#). *Linguistics*, 53(3):479–524.
- John Heritage. 2012. [Epistemics in action: Action formation and territories of knowledge](#). *Research on Language & Social Interaction*, 45(1):1–29.
- Peter Koch and Wulf Oesterreicher. 1985. Sprache der Nähe - Sprache der Distanz. *Romanistisches Jahrbuch*, 36:15–43.
- Gilbert Lazard. 1999. Mirativity, evidentiality, mediativity or other? *Linguistic Typology*, 3:91–109.
- Natalia Levshina. 2017. [Online film subtitles as a corpus: an n-gram approach](#). *Corpora*, 12(3):311–338.
- Max Planck Institute for Evolutionary Anthropology. 2015. [The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses](#).
- David Meyer, Achim Zeileis, and Kurt Hornik. 2006. [The strucplot framework: Visualizing multi-way contingency tables with vcd](#). *J. Stat. Softw.*, 17(3).
- Estelle Moline. 2009. Présentation: Panorama des emplois de *comment* en français contemporain. *Travaux de linguistique*, 58(1):7–17.
- Claudine Olivier. 1985. L’art et la manière: *comment* dans les stratégies discursives. *Langages*, 20:71–98.
- Matthew Purver. 2004. [The theory and use of clarification requests in dialogue](#). Ph.D. thesis, King’s College Univ. of London.
- R Core Team. 2025. [R: A language and environment for statistical computing](#).
- K. Johan Sæbø. 2016. “How” questions and the manner-method distinction. *Synthese*, 193(10):3169–3194.
- Lila San Roque, Simeon Floyd, and Elisabeth Norcliffe. 2018. Egophoricity: An introduction. In Simeon Floyd, Elisabeth Norcliffe, and Lila San Roque, editors, *Egophoricity*, pages 1–78. John Benjamins, Amsterdam.
- John Searle. 1969. *Speech acts*. Cambridge Univ. Press, Cambridge.
- Anna Siewierska. 2004. *Person*. Cambridge Univ. Press, Cambridge.
- Nicolas Tournadre. 2008. Arguments against the concept of ‘conjunct’/‘disjunct’ in Tibetan. In Brigitte Huber, Marianne Volkart, and Paul. Widmer, editors, *Chomolangma, Demawend und Kasbek*, pages 281–308. Int. Inst. for Tibetan and Buddhist Stud., Halle.
- Lucia M. Tovenà. 2023. [Asking about the reason for an effect, and its consequences for the analysis of wh-interrogatives](#). In Łukasz Jędrzejowski and Carla Umbach, editors, *Non-interrogative subordinate wh-clauses*, pages 381–409. Oxford Univ. Press, Oxford.
- Danièle Van de Velde. 2009. *Comment*, manières d’être et manières de faire. *Travaux de linguistique*, 58(1):39–61.
- Manuel Widmer and Fernando Zúñiga. 2017. [Egophoricity, Involvement, and Semantic Roles in Tibeto-Burman Languages](#). *Open Linguistics*, 3(1):419–441.
- Achim Zeileis, David Meyer, and Kurt Hornik. 2007. [Residual-based shadings for visualizing \(conditional\) independence](#). *J. Comput. Graph. Stat.*, 16(3):507–525.

A Glosses

List of abbreviations for glossing according to [Max Planck Institute for Evolutionary Anthropology \(2015\)](#): 1 = first person, 2 = second person, 3 = third person, AUX = auxiliary, COND = conditional, INF = infinitive, IPFV = imperfective, NEG = negation, PL = plural, PRS = present, PST = past, PTCP = participle, REFL = reflexive, SBJV = subjunctive, SG = singular.³

- (1a) Comment va l'enfant ?
how go.3SG.PRS the-child
- (2a) Comment vous faites quand vous
how you.2PL do.2PL.PRS when you
êtes toute seule ?
be.2PL.PRS all alone
- (3a) Comment tu peux imaginer
how you.2SG can.2SG.PRS imagine.INF
que je puisse être assez fou
that I can.1SG.SBJV be.INF enough crazy
pour avoir fait ça ?!
to have.INF do.PST.PTCP that
- (4) J'étais comment ?
I-was.1SG.IPFV how
- (5) Comment je fais si
how I do.1SG.PRS if
y'a rien ici ?
there-have.3SG.PRES nothing here
- (6) Comment peux-tu être aussi
how can.2SG.PRS you.2SG be.INF so
naïf ?
naïve
- (7b) Comment tu sais qu'elle
how you.2SG know.2SG.PRS that-she
boit ?
drink.3SG.PRS
- (8) Comment ils ont pu
how they.3PL AUX.3PL.PRS can.PTCP
lui faire ça ?
him.DAT do.INF that
- (9) Comment j'ai pu ne
how I-AUX.1SG.PRS can.PTCP NEG
rien voir... comment j'ai
nothing see.INF how I-AUX.1SG.PRS
pu me tromper comme ça ?
can.PTCP REFL err.INF like that
- (10) Comment je pourrais ne pas être
how I can.1SG.COND NEG be.INF
jaloux ?
jealous
- (11) Comment je paie la commande ?
how I pay.1SG.PRS the order

B Annotations

Table 5: Confusion matrix of control annotations and corresponding subset of final annotations

		Control				Sum
		man	mea	meth	rea	
Final	man	152	7	6	3	168
	mea	1	108	1	3	113
	meth	6	14	89	3	112
	rea	2	6	6	47	61
	Sum	161	135	102	56	454

³We thank our reviewers for suggesting this step.

C Figures

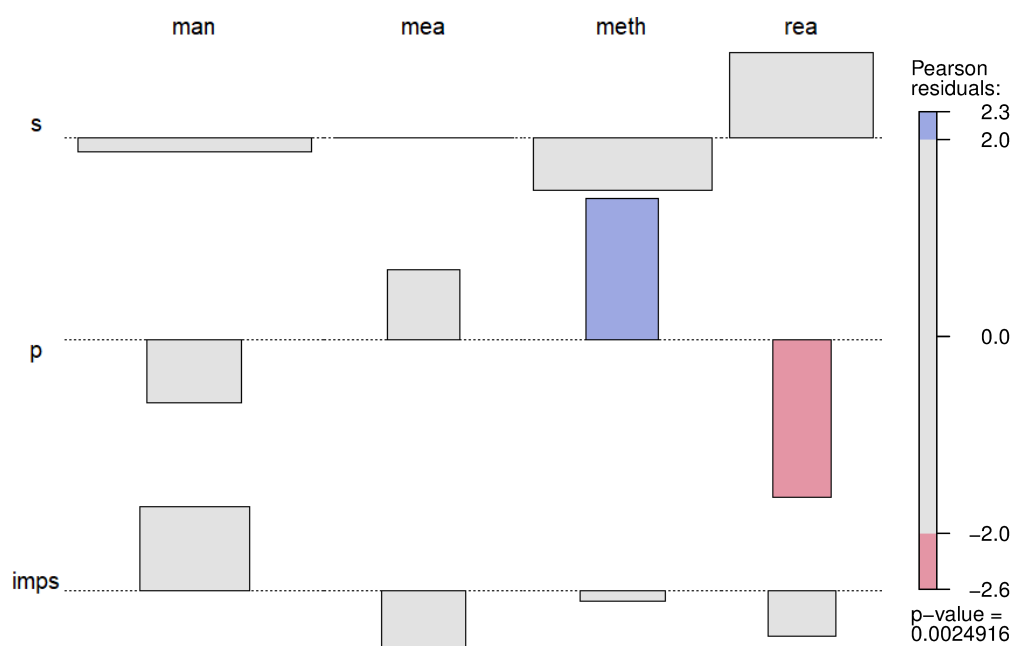


Figure 2: Association plot based on a χ^2 test on Table 3 (number categories by reading).

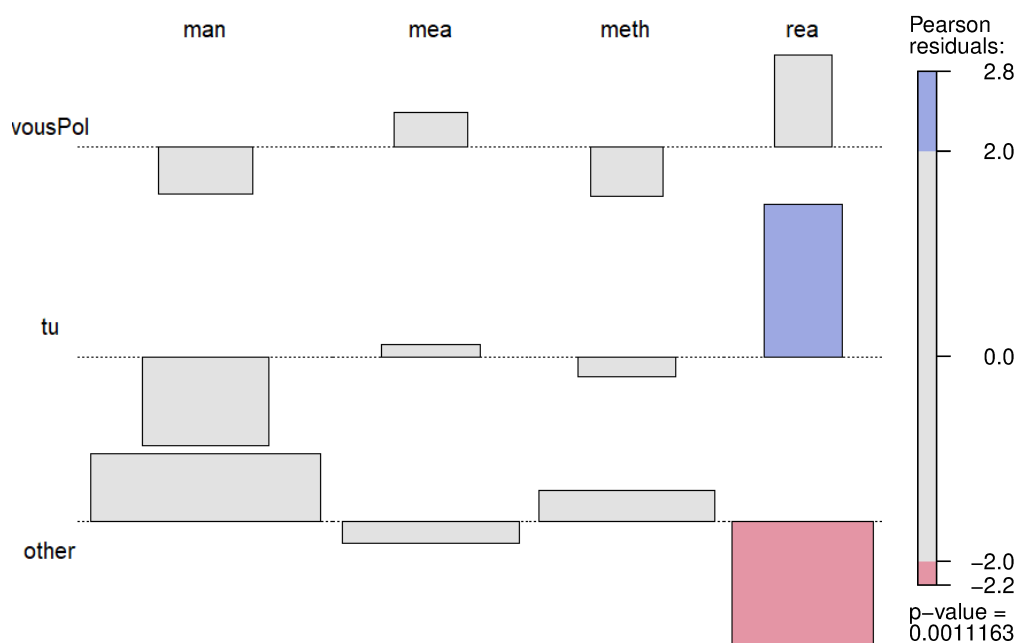


Figure 3: Association plot based on a χ^2 test on Table 4 (second-person address by reading).