

Covert movement in English probing wh-questions

An Nguyen & Géraldine Legendre *

Abstract. Besides fronted information-seeking questions, English also allows for two types of wh-in-situ ones: echo questions, which are used to request a repetition or a clarification of a previous utterance, and probing questions, which are often used in quiz shows, classroom settings, and child-directed speech to “prompt” the addressee for an answer. An acceptability judgment task shows that PQs with multiple wh-phrases get a significantly lower acceptability score than echo questions with multiple wh-phrases despite their similarity in surface structure, which suggests a syntactic difference below the surface. Independent syntactic evidence confirms the result and further suggests that while echo questions involve no syntactic movement (Dayal, 1996), probing questions involve covert wh-movement.

Keywords. wh-question; echo question; probing question; wh-in-situ; wh-movement; covert movement

1. Introduction. Besides fronted information-seeking questions, English also allows for two types of wh-in-situ ones: echo questions (EQs) and probing questions (PQs). EQs are used to request a repetition or a clarification. PQs on the other hand are often used to “prompt” the addressee for an answer, which often contains new information. They appear frequently in quiz shows (Comyn, 2013), classroom setting, and child-directed speech. A corpus analysis of almost 10,000 wh-questions extracted from CHILDES reveals that wh-in-situ take up 16.1% of all main clause questions, and PQs take up 65% of all full sentential in-situ wh-questions (Nguyen, 2019). Examples of all three types of questions are shown below. We adopt the convention of representing wh-phrases in an echo question in caps and those of a probing question in lowercase italic.

- (1) (Information-seeking question)
Where are you going?
- (2) (Echo question)
A: Mary gave a book to John. B: Mary gave WHAT to John?
- (3) a. (Probing question, example from the HSLLD corpus)
On Thursday he ate through 4 strawberries, but he was still hungry. On Friday he ate through 5 oranges, but he was still *what*?
b. (Probing question, example from COCA corpus)
I need to know about displacements. They have a *what*?

EQs and PQs differ from each other in various linguistic aspects. For example, while EQs can only occur as a reaction to a prior utterance (Banfield, 1982; Horn, 1989), PQs can be the first utterance in discourse. EQs have a rising pitch and heavy stress on the wh-word (Pierrehumbert, 1980; Bolinger, 1987; Artstein, 2002), but PQs have a flat/falling pitch. Most notably, despite the similarity in surface (in situ) structure, EQs and PQs are different syntactically. EQs have been claimed to have no covert wh-movement (Dayal, 1996; Artstein, 2002). We will compare them and show on the basis of experimentally collected data that PQs, on the other

*We would like to thank anonymous conference reviewers and the audiences at LSA 2020 for their useful feedback. Authors: An Nguyen, Johns Hopkins University (an.nguyen@jhu.edu) & Géraldine Legendre, Johns Hopkins University (legendre@jhu.edu).

hand, involve covert movement, as they are sensitive to locality constraints and can license Antecedent-Contained Deletion as well as parasitic gaps.

2. Locality constraints. Fronted wh-words in English typically give rise to syntactic locality effects. For example, as shown in (4), an object or adjunct wh-phrase cannot move across a subject wh-phrase. In (4a), the overt movement of the lower (object) wh-phrase across the higher (subject) wh-phrase in a multiple wh-question violates the Superiority Condition (Chomsky, 1973). (4b) contains a wh-island, as an embedded wh-phrase overtly moves across another embedded wh-phrase (Ross, 1967).

- (4) a. *What did who give to Mary?
b. *Which book does Tom know where Jill borrowed from?

In contrast to (4), in-situ EQs containing multiple wh-words in English are grammatical hence not subject to these locality constraints (e.g., Kuno and Robinson, 1972; Artstein, 2002; Sobin, 2010; Beck and Reis, 2018). For example, (5a) is grammatical despite violating the Coordinate Structure Constraint (Ross, 1967), and (5b) is grammatical in contrast to (4a) despite violating the Superiority Condition. This has been interpreted as solid evidence that in-situ EQs do not involve covert movement at LF. The wh-phrase in EQs is genuinely in situ, hence they are not constrained by locality conditions.

- (5) a. A: I know Mary bought cakes and candles
B: You know WHO bought WHAT and candles?
b. A: What did John give to Mary?
B: What did WHO give to Mary?

Yet, PQs parallel to (5a) sounds much less acceptable:

- (6) a. ?*Now class, tell me *who* invented *what* and the light bulb?
b. ?* Now class, tell me *who* invented the light bulb and *what*?

To gain a better understanding of their respective syntactic structures, we conducted an acceptability judgment task on EQs and PQs with multiple wh-words. If in-situ PQs with multiple wh-words violating locality constraints are rated with a high acceptability score like in-situ EQs, then it would indicate that PQs are not subject to locality constraints and are syntactically indistinguishable from EQs. If, however, they receive a low score then their structure arguably involves covert movement at LF.

2.1. A COMPARATIVE EXPERIMENTAL INVESTIGATION

2.1.1. METHOD

Participants

Participants included 138 native English speakers ($M = 19.42$ years, $SD = 1.27$; 47 males), all of whom were undergraduate students at Johns Hopkins University. Data of two participants were excluded from analysis because they either gave unusually high scores for unacceptable controls, unusually low scores for acceptable controls, or both. This left us with 136 data points.

Materials and Procedure

Participants were given an acceptability judgment task with 38 items. Eleven EQs with two wh-words were matched closely in surface structure with eleven PQs with two wh-words. Six PQs with one wh-word were included to confirm that a) PQs are generally well-accepted and b) the discrepancy between EQs with two wh-words and PQs with two wh-words (if any) is not

influenced by the difference in pragmatics between the two types of questions. EQs with one wh-word were not assessed as they are generally assumed to be grammatical (Blakemore, 1994; Sobin 2010). Participants were asked to rate the acceptability of each question on a 7-point Likert scale, with 1 being “least acceptable” and 7 being “most acceptable”. The sentences were presented to the participants both in text and audio forms. Participants had to play the audio file at least once before being able to move onto the next sentence.

All the recordings of the test items were carefully recorded by a female native English speaker to sound natural. We also extracted 24 PQs and 24 EQs from three audio corpora, HSLLD (Dickinson & Tabors, 2001), VanHouten (Van Houten, 1986), and Weist (Weist & Zev-enbergen, 2008), and conducted an acoustical analysis to ensure that the prosodic properties of the recorded test items were identical to those of wh-words in naturally occurring speech. The analysis yielded no significant difference between the recorded items and spontaneous speech.

For each question, a brief context along with the text transcription of the recording was displayed. The task consisted of 38 test items. The design of the test items is shown in Table 1.

Category	# of items	Example
EQs with two wh-words	11	[Context: A is complaining to his friend.] A: I went to the bar to get the company’s report that I forgot. B: <i>You went where to get what that you forgot?</i>
PQs with two wh-words	11	[Context: The teacher is reviewing the materials from last class.] Student: I don’t know how to answer this question. Instructor: Think about the last lecture. <i>Columbus went where to find what that was important?</i>
PQs with one wh-word	6	[Context: The teacher is asking the class a question.] Student: I don’t know how to answer this question. Teacher: Think about the last lecture. <i>In 1879 Edison invented what?</i>
Ungrammatical sentences served as both fillers and controls	5	[Context: Kid sees a boy in the park and asks Mom about him.] <i>What is the boys over there wears red socks is doing?</i>
Grammatical sentences served as both fillers and controls	5	[Context: Two classmates discuss their group project.] <i>What do you think we would need to do to get this project finished?</i>

Table 1. Test items in the acceptability judgment task.

2.1.2. RESULTS. The results show that participants tended to find PQs with two wh-phrases less acceptable than EQs. Although the raw acceptability scores of PQs with one wh-phrase and EQs with two wh-phrases ($M = 5.15$ and $M = 4.72$, respectively) are moderate on a 7-point scale, it is important to note that the mean score of the five grammatical controls is only 5.9. This is most likely due to a mild response style on the part of the participants that prompted them to avoid the extremes of the scale in favor of mid-range responses (Hurley, 1998).

PQs with two wh-phrases were given a much lower score ($M = 3.58$). A nonparametric Mann-Whitney U test indicates that the difference between the acceptability score of EQs with two wh-phrases and PQs with two wh-phrases was significant ($W = 1521700$, p -value $< .0001$). This result still remains significant after applying a z-normalization method across participants to

take individual differences into account. Figure 1 summarizes the results of the acceptability judgment task. Probing_1 represents PQs with one wh-phrase while Probing_2 represents PQS with 2 wh-phrases.

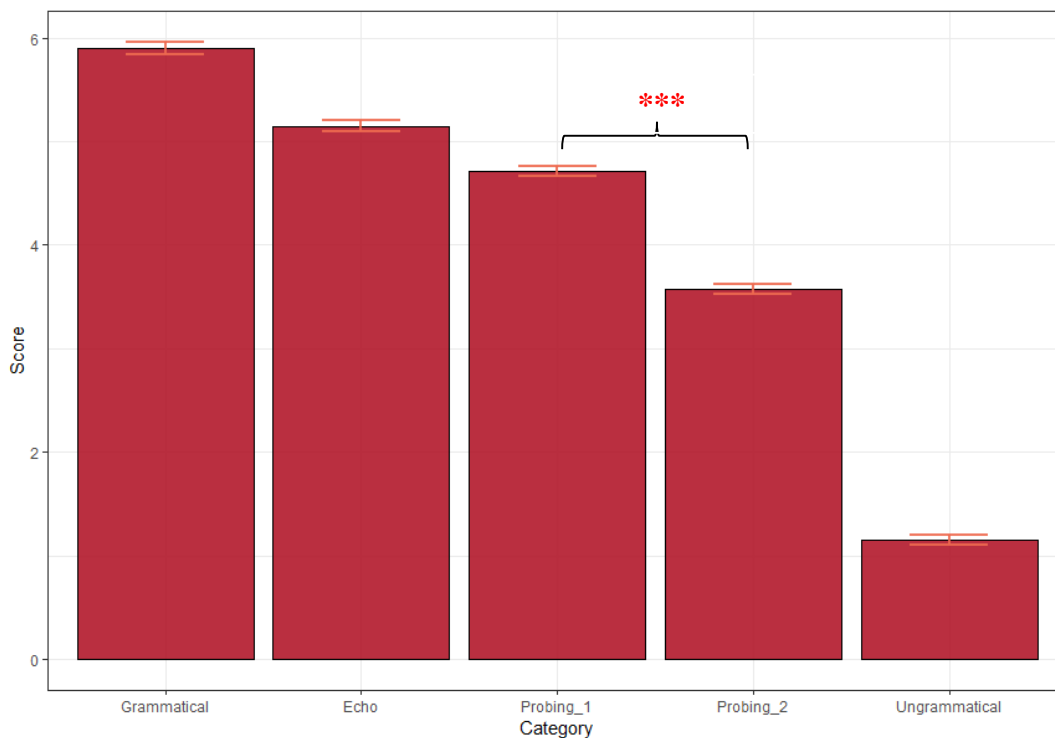


Figure 1. Mean score of acceptability across 5 categories of questions.

The data we obtained experimentally show that in-situ EQs with two wh-phrases are rated with a significantly higher acceptability score than in-situ PQs with two wh-phrases. This, in turn, suggests that in-situ EQs are not subject to locality constraints, which is in line with previous literature on EQs in English (Kuno and Robinson, 1972; Artstein, 2002; Sobin, 2010; Beck and Reis, 2018). We follow Dayal (1996) in endorsing an analysis whereby EQs do not undergo covert wh-movement at LF, which explains their immunity from locality effects. At the same time, the novel result that PQs with two wh-phrases receive a significantly lower score than EQs with two wh-phrases, despite their surface similarity, suggests that in-situ PQs are in fact subject to locality constraints. Since overt and covert A-bar movement have been construed as subject to the same locality constraints (Choe, 1987; Nishigauchi, 1990), we propose that PQs, in contrast to EQs, do involve covert wh-movement at LF. Independent syntactic evidence is presented next, which confirms this conclusion.

3. Antecedent-Contained Deletion licensing. If a phrase can contain an instance of Antecedent-Contained Deletion (ACD), then the phrase undergoes covert movement (Pesetsky, 2000). An example of ACD is presented in (7):

(7) John read [every book that Mary did [VP Δ]].

The sentence above is interpreted as in (8):

(8) John read [every book that Mary [VP read t]].

The elided VP in (7) is interpreted as the VP [*read t*], hence it must have an antecedent of the form [*read t*] in the linguistic context. This is resolved by covert movement of the DP containing [*VP read t*] to a VP-external position (specAgrOP):

(9) [AgrOP Every book that Mary did [*VP read-t*]] [John [*VP read t*]].

Multiple wh-questions in English, for example, support ACD, which is evidence that the in situ object wh-phrase containing the elided VP undergoes covert wh-movement (Pesetsky, 2000):

(10) Which girl invited which student that John did [*VP Δ*]?

If PQs involve covert movement, they too should be able to license ACD. As corpus data shows that parents frequently switch to a PQ after children fail to answer a fronted information-seeking question (Nguyen, 2019), let's consider the following scenario: A mother is teaching her son the concept of equivalence. She asks her son: "John knew that Bill read a certain number of books, and he also thought that Fred read the same amount of books as Bill. *How many books did John think that Fred read?*". When the son fails to answer the question, she asks it again as a PQ, adding a hint about the books that Bill read:

(11) John thought that Fred read *how many* books that Bill did?

Eight out of 8 native speakers of English we consulted report that the question above is grammatical. This, in turn, is evidence that covert wh-movement is involved:

(12) [AgrOP how many books that Bill did [*VP read-t*]] [John thought that [Fred [*VP read t*]]]?

Note that we do not consider the possibility of EQs licensing ACD because it is pragmatically infelicitous to ask such questions. The EQs in (13), serving as a response to the statement "John read every book that Mary did", sound unnatural according to the 8 native speakers we consulted.

(13) a. #John read HOW MANY book that Mary did?
b. #John read WHAT book that Mary did?

4. Parasitic gaps. Parasitic gaps (PG) are licensed by a gap left by A-bar movement (Engdahl, 1983), which includes traces of wh-movement (Chomsky, 1986). (14a) illustrates a PG licensed by wh-movement, while (14b) shows that PG is not possible when there is no A-bar movement.

(14) a. Which article_i did Mary file t_i without proof-reading *pg*?
b. *Mary filed the article without proof-reading *pg*.

Assuming covert movement works in a similar way to overt movement, we expect that PQs will license PG.

Consider the following scenario: All students in a research methods class have to read about Mary, a careless researcher. One time, Mary forgot to proofread a paper before submitting it for publication. All the students who read the story should know which paper this was. During a discussion, the teacher asks the class:

(15) ?So, Mary filed *which* article without proof-reading *pg*?

Seven out of 8 native speakers we consulted report that in a proper context, (15) is relatively acceptable. All eight of them, however, agree that (15) is much better than (14b), though not as good as (14a). Assuming that (15) is indeed grammatical, this would provide additional evidence

that PQs involve covert movement. We suggest that the relative difference in acceptability (15 vs. 14a) may be tied to the overt/covert distinction, in particular, the absence of an s-structure wh-trace licensing PG in (15).

EQs cannot be considered in this case. In (16), the previous utterance (16A) is ungrammatical, making it impossible to have an EQ following it. While (17B') with a pronominal copy is grammatical, the EQ in (17B) is rejected because it violates the restriction that EQs must reflect the structure of the preceding utterance.

- (16) A: *Mary filed the X article without proof-reading *pg*.
B: Mary filed WHICH ARTICLE without proof-reading *pg*?
- (17) A: Mary file the X article without proof-reading it.
B: *Mary filed WHICH ARTICLE without proof-reading *pg*?
B': Mary filed WHICH ARTICLE without proof-reading *it*?

5. Conclusion. Both experimental evidence on EQs and PQs with multiple wh-phrases and independent syntactic evidence confirm the syntactic difference between PQs and EQs. PQs are sensitive to locality constraints and can license ACD and parasitic gaps. EQs, on the other hand, are not subject to locality constraints and cannot license ACD and parasitic gaps. Hence, PQs involve covert wh-movement while EQs involve no movement. This result poses challenges to existing analyses of wh-movement in English. Furthermore, investigating the difference in movement types across different categories of wh-questions (i.e., overt movement in information-seeking question, covert movement in PQs, and no movement in EQs) would likely lead to more insights into movement-driving features in questions.

References

- Artstein, Ron. 2002. A focus semantics for echo questions. In Ágnes Bende-Farkas and Arndt Riestler (eds.), *Workshop on Information Structure in Context*. 98-107. IMS, University of Stuttgart.
- Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston, MA: Routledge & Kegan Paul.
- Beck, Sigrid & Marga Reis. 2018. On the form and interpretation of echo *wh*-questions. *Journal of Semantics* 35(3). 369–408. <https://doi.org/10.1093/jos/ffy002>.
- Blakemore, Diane. 1994. Echo question: A pragmatic account. *Lingua* 94 (4). 197-211. [https://doi.org/10.1016/0024-3841\(94\)90009-4](https://doi.org/10.1016/0024-3841(94)90009-4).
- Bolinger, Dwight. 1987. *Intonation and its parts: Melody in language*. Stanford: Stanford University Press.
- Choe, Jae-Wong. 1987. LF wh-movement: A case of pied piping. *Linguistic Inquiry* 18. 348–353.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*. 232–286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.
- Comyn, Marie. 2013. *Wh-in situ is accompanied by which formal features?* Ghent University dissertation.
- Dayal, Veneeta. 1996. Locality in wh-quantification: Questions and relative clauses in Hindi. *Studies in Linguistics and Philosophy* 62. 125. <https://doi.org/10.1007/978-94-011-4808-5>.

- Dickinson, David, & Patton Tabors. (Eds.) (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore: Paul Brookes Publishing.
- Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy*. 6. 5–34. <https://doi.org/10.1007/BF00868088>.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago; Chicago University Press.
- Hurley, John. 1998. Timidity as a response style to psychological questionnaires. *Journal of Psychology* 132. 201–210. <https://doi.org/10.1080/00223989809599159>
- Kuno, Susumu & Jane Robinson. 1972. Multiple wh questions. *Linguistic Inquiry* 3(4). 463-487.
- Nguyen, An. 2019. *The acquisition of wh-questions: a case of probing questions*. Johns Hopkins University Master thesis.
- Nishigauchi, Taisuke. 1990. Quantification in the theory of grammar. *Studies in Linguistics and Philosophy* 37. 116-117. https://doi.org/10.1007/978-94-009-1972-3_4.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge Mass: MIT Press. <https://doi.org/10.7551/mitpress/5365.001.0001>.
- Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Massachusetts Institute of Technology dissertation.
- Ross, John R. 1967. *Constraints on variables in syntax*. Massachusetts Institute of Technology dissertation.
- Sobin, Nicholas. 2010. Echo questions in the minimalist program. *Linguistic Inquiry* 41(1). 131–148. <https://doi.org/10.1162/ling.2010.41.1.131>.
- Van Houten, Lori. 1986. *Role of maternal input in the acquisition process: The communicative strategies of adolescent and older mothers with their language learning children*. Paper presented at the Boston University Conference on Language Development, Boston.
- Weist, Richard & Andrea Zevenbergen. 2008. Autobiographical memory and past time reference. *Language Learning and Development* 4(4). 291 – 308.