1 Chung (2010) on *wh*-movement in Chamorro
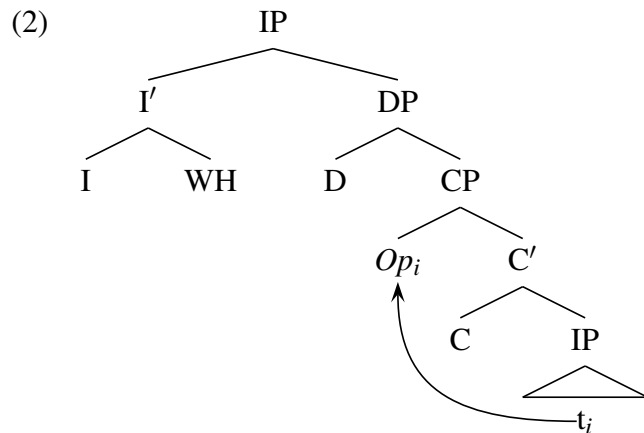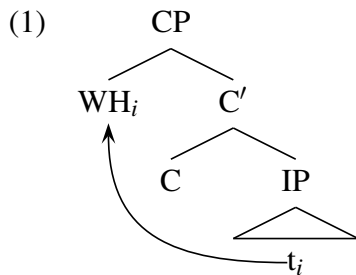
1.1 The issue

> Greenberg's (1963) Universal 12:
> If a language has dominant order VSO in declarative sentences, it always puts interrogative words or phrases first in interrogative word questions; if it has dominant order SOV in declarative sentences, there is never such an invariant rule.

There are at least two possible approaches to constituent questions (a.k.a. interrogative word questions, a.k.a. *wh*-questions) in VSO languages.

- They are formed with *wh*-movement of the kind familiar from SVO languages (1).

- In VSO languages, *wh*-words are "verbs" and so a *wh*-question has a form abstractly more like "(That which) John bought is what?"—the generalization that the *wh*-word comes first is just a specific case of the generalization that the verb comes first (2).

(1)

```
        CP
       /  \
    WH_i   C'
     ↑    /  \
     |   C    IP
     |        /\
     +------t_i
```

(2)

```
             IP
           /    \
         I'      DP
        /  \    /  \
       I   WH  D    CP
                   /   \
                 Op_i   C'
                  ↑    /  \
                  |   C    IP
                  |        /\
                  +------t_i
```

1.2 Building the argument

To build the argument for (1) over (2) (or for that matter just to do the research to convince oneself), the strategy is to think about anything that the choice between them might interact with. What differentiates the two hypotheses?

They have in common the basic fact that the *wh*-phrase is expected to come first. But in (1), the *wh*-phrase has moved to where it is, and in (2), the *wh*-phrase simply starts out first. The *Op* that has moved in (2) is part of what makes a relative clause, but there is basically a kind of *wh*-movement happening in both—the question is really whether it is the *wh*-phrase itself that has moved, or whether it is this silent *Op* that has moved. And consequently, whether the *wh*-phrase is itself the main predicate, or is structurally higher than the main predicate.

If you want to test for a movement relation, the first thing you'd probably think of to check is whether there are **island constraints**—if you find that an island boundary between the putatively moved item and

its trace renders the sentence ungrammatical, then that constitutes support for a movement analysis. Here, though, that is not very useful, since there is a movement in both cases and nowhere one could really place an island boundary that would break the path of movement in one but not the other.

Apart from that, there's a certain amount of fishing around—what do we know independently about the language that would treat the structures in (1) and (2) differently?

Chung (2010) looks at a couple of different types of phenomena.

- In (2), the *wh*-phrase is a predicate. So any independent constraints on what kinds of predicates there are should also apply to *wh*-phrases under analysis (2) (but not under (1)).

- In (1), the *wh*-phrase was one the verb's original arguments, moved from a lower position in the structure. Thus, constraints imposed by the verb on its arguments should hold of the *wh*-phrase (but not necessarily under (2)). Also, constraints on *wh*-phrases from Binding Theory that affect the pre-movement position should limit the form of the *wh*-phrase (but again, not under (2)).

- If Chamorro is to be analyzed as in (2), we expect it to behave like other languages which have been argued (more persuasively, I guess) to have a structure like (2).

- In (1), the *wh*-phrase is higher in the structure than the predicate, while in (2) the *wh*-phrase *is* the predicate. Predicates appear not to c-command the subject—so, using the closely related focus construction, we might explore whether a focused object c-commands the subject or not. Assuming what is true of focus is also true of *wh*-questions.

- Constraints on subjects of DP predicates should also hold of subjects of DP *wh*-words under (2).

1.3 Chamorro word order and case

The unmarked order is Predicate—Subject—Complements. Case marking precedes the DP, and is either "unmarked" (UNM), "oblique" (OBL), and "local" (LOC). I will use WHOBL to mean "*wh*-agreement for oblique," WHNOM "*wh*-agreement for nominative/subject," and WHOBJ "*wh*-agreement for object."

(3) Ha-konfitma i   kotti i   intensión i   Covenant Agreement.   (V)
    AGR-confirm the court the intension the Covenant Agreement
    'The court confirmed the intention of the Covenant Agreement.'

(4) Hägas   ha-läknus   ennao siha na planu si   Speaker Benigno R. Fitial.   (Adj)
    long.ago AGR-present that   PL   L plan UNM Speaker Benigno R. Fitial
    'Speaker Benigno R. Fitial presented those plans long ago.'

(5) Ti médiku esti siha na siñoris.   (NP)
    not doctor this PL   L gentlemen
    '[The public is fortunate that] these gentlemen are not doctors.'

(6) Ginin i   asagua-ña   gi   fine'nena dos hana-ña.   (PP)
    from the spouse-AGR LOC first   two daughter-AGR
    'Two of his daughters were from his first wife.'

2 The arguments

2.1 An impossible type of predicate

Independent fact: DPs in the locative case (was this what Chung called "local" earlier?) cannot be the predicate of a clause.

(7)   a.   Ginin as   Juan i   katta.
           from OBL Juan the letter
           'The letter was from Juan.'

      b.   * Gi   petta i   ispehus.
             LOC door the mirror
             ('The mirror is on the door.')

      c.   * Gi   hilu' lamäsa magagu-mu.
             LOC top   table clothes-AGR
             ('Your clothes are on top of the table.')

- (2) predicts that DP *wh*-words in locative case should also be impossible.

- (1) doesn't—all things being equal, they should be fine.

(8)   a.   Gi   manu ni   man-ma'a'ñao i   famagu'un mañ-aga   _ ?
           LOC where? COMP AGR-afraid   the children   INFIN.AGR-stay
           'Where are the children afraid to stay?'

      b.   Giya hayi nai   ha-dipépendi   gui'   si   Juan _ ?
           LOC who? COMP AGR-depend.PROG himself UNM Juan
           'Who does Juan depend on?'

> Not mentioned, but the relevance of (8a) is ensured by the fact that *manu* 'where?' is case marked, indicating that it is in fact a DP. Otherwise, it might have been some kind of adverb/adjunct.
> Also, note that we don't have an explanation for *why* the locative DPs are not allowed as predicates. We don't really need to know why for the purpose of this argument—so long as the generalization itself is basically correct.

2.2 Connectivity in selectional restrictions

This is an argument about whether there is evidence that the *wh*-phrase was once a complement of the verb—if so, it is an argument in favor of (1).

To get us started, a phenomenon in English that we will compare to Chamorro.

(9)   a.   They went to Tinian.

      b.   * They went on Tinian.

      c.   * They stayed to Tinian.

      d.   They stayed on Tinian.

In English "*wh*-clefts" like (10), this selection seems to be preserved. Note here that the analysis is far from clear—but the idea is that the focus (*to Tinian* or *on Tinian*) must have undergone movement. The fact that the selection constraints are still in effect tells us that the focus was once a complement of the verb. How *where* and *was* fit in here is not clear—but it is a topic that's beyond what we can really explore right now.

(10) a. Where they went was to Tinian.
   b. * Where they went was on Tinian.
   c. * Where stayed went was to Tinian.
   d. Where stayed went was on Tinian.

If we bury the *wh*-clause inside a headed relative (*the island where they went/stayed*) then the choice of verb doesn't have any effect.

(11) a. The island where they went was Tinian.
   b. * The island where they went was to Tinian.
   c. * The island where they went was on Tinian.
   d. The island where they stayed was Tinian.
   e. * The island where they stayed was to Tinian.
   f. * The island where they stayed was on Tinian.

In Chamorro, the idea is that there are similar selection relations between 'take' ('to') and 'put' ('on').

- (1) predicts that the selection constraints should still be in force for the *wh*-phrase.

- (2) predicts (sort of) that the selection constraints should not be able to "reach" the *wh*-phrase.

(12) a. Pära manu guätu nai ma-konni' si Miguel _ ?
       to where? over.there COMP AGR-take UNM Miguel
       'To where did they take Miguel?'
   b. Manu guätu nai ma-po'lu i fmagu'on-ta _ ?
       where? over.there COMP AGR-put the children-AGR
       'Where did they leave our children?'
   c. * Pära manu guätu nai ma-po'lu i fmagu'on-ta _ ?
       to where? over.there COMP AGR-put the children-AGR
       ('To where did they leave our children?')

A critical eye: We were not given the examples in (9) to build up the case, nor were we given their analogs in Chamorro. Also, unless there is some reason that it is not possible, it is kind of important to see something like (11) in Chamorro relative clauses for comparison—if Chamorro headless relatives *also* exhibit this kind of selection, then we can't conclude anything from this argument.

Maybe more importantly, the rough state of our understanding of the English *wh*-clefts makes it difficult to really compare them directly to the Chamorro analogs. Perhaps, if we had the Chamorro analogs, the analysis would be clearer. But the actual source of the connectivity effects in (10) is unclear—direct selection is one possibility, but perhaps there is also something slightly more indirect going on. It would certainly be reasonable to suppose that *where* is actually the thing that has undergone *wh*-movement, but then how are the selectional effects communicated? All in all, I find this argument to be kind of non-compelling.

2.3 Connectivity in antecedent-pronoun relations

This is basically an argument from Principle C, but Chung claims that there is an additional requirement that Chamorro is subject to.

Recall that Principle C says that a name ("r-expression") cannot be c-commanded by something that is coindexed with it. Hence (13) is bad.

(13)   * He$_i$ saw John$_i$ in the mirror.

Principle C is all about c-command, so if *he* doesn't c-command *John*, then there's no problem.

(14)   John$_i$ gave [the book he$_i$ read] five stars.

(15)   [The book he$_i$ read] gave John$_i$ a headache.

Not so in Chamorro, though—not only must Principle C be obeyed, but the r-expression has to come *before* any pronoun that it is coindexed with. So, (15) would not have been good in Chamorro. But (14) would have been.

But, wait—there's more. The question we're trying to answer is whether the *wh*-phrase in (1) was once lower in the structure. It turns out that Principle C can tell us this, because it seems that Principle C must be satisfied when the *wh*-phrase is in its original position (16).

(16)   * [Which song about John$_i$] did he$_i$ like?

(17)   [Which song about him$_i$] did John$_i$ like?

So, now we have the pieces in place—in particular, we're looking for something like (16) in Chamorro. If the *wh*-phrase in (16) was never c-commanded by *he*, then there should be no Principle C violation. More specifically,

- (1) predicts that the analog to (16) should be out, due to the Principle C violation under reconstruction.

- (2) doesn't—the *wh*-phrase was never c-commanded by the r-expression (and *Op* is not coindexed with the name).

- (2) predicts that the analog to (17) should be out, because the coindexed pronoun precedes the name.

- (1) doesn't necessarily—it could go either way, depending on what really is behind the constraint that the pronoun follows the name. In its pre-movement position, the pronoun "followed" the name, so if the analog to (17) is ok, that must have been sufficient.

There is also a mild complication in the fact that the pronouns in Chamorro can be silent (written as *pro*), but that shouldn't change anything really.

Without further ado, here's what we have: The analog to (16) is (18), and the analog to (17) is (19).

(18)   Kuantu   gi   tumobit Juan$_i$ esta   ha-fa'gasi   *pro$_{j,*i}$* _ ?
       how.many? LOC car   Juan already WHOBJ.AGR-wash ↑
       'How many of Juan's$_i$ cars did he$_{j,*i}$ wash?'

(19)   Kuantu   gi   tumobet-ña *pro$_i$* esta   ha-fa'gasi   si   Juan$_i$ _ ?
       how.many? LOC car-AGR   ↑   already WHOBJ.AGR-wash UNM Juan
       'How many of his$_i$ cars did Juan$_i$ wash?'

Conclusion: The predictions of (1) were met, not those of (2).

Based on this, the requirement that a pronoun follows its antecedent is a bit weird—because it doesn't really have anything to do with pronunciation, apparently, yet it would otherwise seem that linear order is all about pronunciation. This makes me suspect that we don't quite know what that extra condition is.

2.4 Sluicing

This is kind of a convoluted argument, but the idea here is that: There's another language (Malagasy) in which it has been argued that *wh*-questions involve headless relative clauses, basically the structure in (2).

Suppose that's true—suppose that Malagasy can be described by (2) more or less. But what about *Chamorro*? Well, if (2) is the right analysis of Chamorro, then we might expect that Chamorro would behave in relevant respects like Malagasy.

One of the arguments in favor of treating Malagasy as a (2)-type language comes from the fact that it does not obey "Merchant's generalization" with respect to "sluicing." So, a brief note about that.

Sluicing is a type of ellipsis that affects the second of two clauses, allowing it to end in a dangling *wh*-word. A simple sluice is shown in (20).

(20) a. We know she bought something, but we don't know what.

b. We know she bought something, but we don't know [$_{CP}$ what [$_{IP}$ ~~she bought~~]]

It looks basically like the IP is simply elided after the *wh*-word moves, and this looks like it might be at least mostly right. The observation Chung refers to as "Merchant's generalization" concerns PP *wh*-phrases like *to whom*. In English, there are two options when asking a question of this kind. Either move *who* by itself and leave *to* behind (21a) ("preposition stranding"), or move them both together (21b) ("pied-piping").

(21) a. Who was she talking to?

b. To whom was she talking?

Not all languages allow preposition stranding—probably most, in fact, don't. For example, French.

(22) a. * Qui est-ce que Pierre l'a offert à?

b. À qui est-ce que Pierre l'a offert?

As part of the argument that the *wh*-word moves to SpecCP, and then IP is elided, Merchant observes that sluicing constructions behave exactly as the overt counterparts do—exactly those languages that allow a preposition to be stranded under *wh*-movement also allow a preposition to be stranded in sluicing.

(23) Pierre l'a offert à quelqu'un, mais je ne sais pas *(à) qui.

(24) Pierre offered it to someone, but I don't know (to) who(m).

The idea (using English) is basically just this:

(25) a. We know she was talking to someone, but we don't know who.

b. We know she was talking to someone, but we don't know [$_{CP}$ who [$_{IP}$ ~~she was talking to~~]]

(26) a. We know she was talking to someone, but we don't know to whom.

b. We know she was talking to someone, but we don't know [$_{CP}$ to whom [$_{IP}$ ~~she was talking~~]]

Makes perfect sense. Now, back to Malagasy. Malagasy disallows P-stranding, but can still strand invisible ones under sluicing. Why? The idea is basically that in Malagasy, it is more like "I forget (the one to whom she was talking is) who"—no stranded preposition. That's the (2) analysis.

So—Chamorro. Chamorro does not allow P-stranding (like Malagasy), and it has sluicing when P-stranding is not involved.

(27) a. Ginin hayi na   un-risibi   i   katta _ ?
          from who? COMP AGR-receive the letter
          'From whom did you receive the letter?'

      b. * Hayi na   un-risibi   i   katta ginin _ ?
          who? COMP AGR-receive the letter from
          ('Who did you receive the letter from?')

(28) Mam-ómoksai   mannuk, lao ti ta-tungu'   hafa   na klasi _
      AGR.AP-raise.PROG chicken but not AGR-know what? L type
      'He's raising chickens, but we don't know what kind.'

On the convoluted trail that we've been winding down, we're at the point that we can isolate the predictions.

- (1) predicts that Chamorro will look like French—no P-stranding even in sluicing.

- (2) predicts that Chamorro will look like Malagasy—P-stranding should be fine in sluicing.

And: (1) wins. Chamorro *cannot* strand a preposition in sluicing.

(29) Si   Joe ha-hunguk i   istoria ginin guahu, lao ti   hu-tungu' …
      UNM Joe AGR-hear the story from me   but not AGR-know …
      'Joe heard the story from me, but I don't know…'

      a. …ginin kuantu   más na taotao _ .
          …from how.many? more L person
          '…from how many others.'

      b. * …kuantu   más na taotao _ .
          …how.many? more L person
          ('…how many others.')

2.5   Negative concord

This argument is again kind of indirect. It isn't actually about *wh*-questions at all. It relies instead on the idea that *focus* constructions have the same kind of syntax as *wh*-questions. So, what we're going to do here is try to see if focus constructions can be analyzed as headless relative clauses, and then generalize from that result to *wh*-questions.

This has to do with "negative concord" which is quite a bit like NPIs. The phenomenon might be familiar from nonstandard English. But it's a pretty common thing (e.g., in Spanish, Italian).

(30)   I ain't givin' nothin' to nobody. ≠ I am giving something to sombody.

(31)   * I'm givin' nothin' to nobody. (=I am giving everybody something)

(32)   Nobody's givin' nothin' to Fred.

(33)   Nobody's givin' sweaters to nobody.

(34)   * Nobody ain't givin' sweaters to Fred.

It's basically the same distribution as NPIs.

(35)   I'm not giving anything to anybody.

(36)   * I'm giving anything to anybody.

(37)   Nobody's giving anything to Fred.

(38)   Nobody's giving sweaters to anybody.

(39)   * Anybody's not giving sweaters to Fred.

The generalization is that the concord forms need to be in the c-command domain of a negative licensor. Above, *ain't* can do that, as can *Nobody* in subject position. Below, we have basically the parallel facts—a negative predicate acts like *ain't* in English: it can license object concord, but not subject concord. The (negative) predicate doesn't c-command the subject, but c-commands the object.

(40)   Ti hu-bisita ni háyiyi ha'.
         not AGR-visit not anyone EMP
         'I didn't visit anyone.'

(41)   * Ti mattu   nigap   ni háfafa   ha'.
         not AGR-arrive yesterday not anything EMP
         ('Anything didn't arrive yesterday.')

If you focus a negative phrase, it will also license negative concord in the remainder (the VP, adjuncts).

(42)   Ni unu lumi'i'   _ si   Dolores ni mánunu ha'.
       not one WHNOM-see   UNM Dolores not anywhere EMP
       '*No one* saw Dolores anywhere.'

And so we arrive at the crucial test. Here's the idea.

- (1), assuming that focus and *wh*-questions are formed in the same way, predicts that focusing a negative object should license subject concord, because it will have *wh*-moved to a position above the subject.

- (2), assuming that focus and *wh*-questions are formed in the same way, predicts that the negative focus *becomes* the predicate, so it should have the same licensing options as the negative predicate did above. In particular, it shouldn't c-command the subject, and so subject concord should be disallowed.

So, we look at a focused negative object and see if it can license negative concord within the subject—and it can. Again, (1) wins.

(43)   Ni háfafa   ha' ma-tätaitai   _ ni unu giya hämi.
       not anything EMP WHNOM.AGR.PASS-read.PROG   not one LOC us
       '*Nothing* had been read by any one of us.'

2.6   A specificity effect

The last argument is a bit like the first one—ultimately, it revolves around a constraint on subjects that requires certain ones to be specific.

One kind of strange thing in this section is that we actually find a bit of evidence *in favor* of an analysis like (2)—but Chung claims that it is only a very small set of examples that must be analyzed in that way, while most cannot.

The phenomenon Chung explores here is a restriction on subjects—when a subject is originally an external argument ("external" here means external to the VP, like the Agents that start in Spec*v*P), it must be "specific." DPs with a quantifier (*meggai na taotao* 'many people') or a null nonspecific article (*lalahi* 'men') don't count as specific, while DPs like names, or with a definite determiner, or an indefinite article or numeral, do count as specific.

(44)   * Mañ-áchalik   lalahi.
       AGR-laugh.PROG men
       ('Men laughed.')

(45)   humyung   patgun.
       AGR.come.out child
       'A child emerged.'

This is true even for headless relatives.

(46)   Áttilung   *(i) *Op* gaigi   _ gi   halum kahita.
       AGR.black the   WHNOM.AGR.be.at LOC inside box
       'The/*A thing that was in the box was black.'

Finally, when a DP is the predicate, it is "individual-level" (meaning that it holds of the subject in a timeless fashion—*John is a doctor*, *John is the president*; cf. *John is hungry*). The main thing is: a subject of a DP predicate is going to count as an external argument.

Let's line up the predictions we can make now.

- (2) predicts that a DP *wh*-word, which is presumed to be serving as a predicate, has an external argument as its subject. This in turn suggests that the headless relative that constitutes the rest of the clause must be specific—something that would be indicated by, e.g., the definite article *i* 'the.' It should not be good without the article.

- (1) does not predict this—as far as (1) goes, there should be no need for things like *i* 'the.'

There *are* in fact sentences that conform to the prediction of (2), in fact. Like (47)—which in fact pretty much must be analyzed as something like (2).

(47)   Hayi ädyu i *Op* pära u-fahani-n   maisa gui' _ present?
       who that the   FUT WHNOM.AGR-buy.for-L self   him   present
       'Who is the one who's going to buy a present for himself?'

So, a structure like (2) is possible, according to this test. "Examples of this sort, however, are clearly in the minority." Most of the time there is no overt article.

(48)   Hafa gaigi   _ gi   tatti-n   petta-n gima'-ñiha.
       what? WHNOM.AGR.be.at   LOC behind-L door-L house-AGR
       '[Not even I know] what is behind the door of their house.'

The point here: *Hafa* 'what' is a DP, its subject on the (2) analysis, must be a headless relative that is an external argument. It should therefore be required to be specific—but yet the sentences where it isn't are grammatical.

The conclusion seems to be that *the majority* of *wh*-questions cannot be analyzed as in (2), even if there are a few that must be.

3   Parting words

So: (a) Chamorro has regular *wh*-movement; (b) it is a VSO language, so Greenberg's universal isn't actually completely right.

References

Chung, Sandra. 2010. Six arguments for *wh*-movement in Chamorro. In Donna B. Gerdts, John C. Moore & Maria Polinsky (eds.) *Hypothesis A/Hypothesis B: Linguistic explorations in honor of David M. Perlmutter*, 91–110. Cambridge, MA: MIT Press.

Denning, Keith & Suzanne Kemmer (eds.) . 1990. *On Language*. Stanford University Press.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.) *Universals of language*, 73–133. Cambridge, MA: MIT Press. Reprinted in Denning & Kemmer (1990), pp. 40–70.