On the Representation of Wh-Words and Foci: Evidence from Mixtec

Abstract: In this paper, I explore the consequences of a cross-linguistic preference for wh-words to move across a more local foci in languages where they compete for a single position. On the basis of novel data from San Martín Peras Mixtec (Otomanguean, Mexico), I show that this general pattern does not only reflect the behavior of exhaustive foci, but in fact applies more generally to all types of foci that move syntactically. I account for this preference by positing that both foci and wh-words move when in the scope of some particle that is sensitive to the alternatives that they introduce. Wh-words are in the scope of a Question Particle (Cable, 2010), and foci are in the scope of other focus sensitive particles. A probe relativized to the features $\text{[ALT]}$ and $\text{[Q]}$ will move a QP instead of a local focus sensitive particle due to a general economy principle—probes “prefer” to be valued by a single goal that can completely satisfy their needs, even if this goal is non-local, rather than by two distinct goals. I argue that this analysis is superior to an alternative that posits distinct heads which probe for wh-words and foci independently and show that the pattern is not reducible to a Focus Intervention Effect.

Key Words: QP Movement, Focus Sensitive Particles, Ā-Probing, Mixtec

1 Introduction

It has long been observed that there are syntactic and interpretive similarities between wh-words and foci. One of these is that they seem to appear in the the same surface position in many languages (Horvath, 1986; Rochemont, 1978, 1986; Chomsky, 1977; Haida, 2007; Aboh, 2007; Rizzi, 1997; É. Kiss, 1998a; Croft, 1990, a.o.). Consider Aghem, for instance, a Bantu language spoken in Cameroon with a default SAuxVO word order (1a). In this language, both wh-words (1b) and focused constituents (1c) surface in a position
immediately following the verb (Watters, 1979; Hyman, 2010).¹

(1) a. Tì-bvù tì-bigâ hà mò zi  kì-be ₄ nè
dogs two PST eat fufu today
‘The two dogs ate fufu today.’

b. À mò zi ndùghà ₄ be ₄ kó nè à
EXPL PST eat who fufu DET today Q
‘Who ate fufu today?’

c. À mò zi tì-bvù tì-bigâ hà be ₄ kó nè
EXPL PST eat dogs two fufu DET today
‘The two dogs ate fufu today.’

The principal aim of this paper is to provide an account of how wh-words and foci are
formally represented in the syntax, with the goal of explaining their surface similarity in a
principled way.

Beyond this surface syntactic similarity, there are other clear indications that wh-
words and foci form a natural class cross-linguistically. First, many languages mark foci
and wh-words morphologically with the same overt particle. This is the case in, for exam-
ple, in Somali (Saeed, 1984), Gungbe (Aboh, 2007), Gùrùntùm (Hartmann and Zimmer-
mann, 2009), and Miyara Yaeyaman (Davis, 2013). In Miyara Yaeyaman, for instance,
both foci and wh-words are followed by the particle du.

(2) a. Taa=du suba tsukur-ee-ru?
who=DU soba make-RES-PRS
‘Who made soba?’

b. Jurie=n=du tsukur-ee-ru
J.=NOM=DU make-RES-PRS
‘Yurie made (soba).’

¹For expository ease, throughout this paper I bold constituents that are interpreted as foci. I use standard
Leipzig Glossing Conventions, with the following additions: ACT= Active Voice, AML= Animal Pronoun,
EMPH= Emphatic Particle, EXPL= Expletive Subject, INT= Interrogative Clause, NEUT= Neutral Pro-
noun, PN= Proper Name Marker, PREV= Preverbal Particle, TAM= Tense, Aspect, Mood Marker, VM= Verb Modifier. Some glosses have been slightly altered from cited sources for the sake of consistency.
Second, in some languages, there are co-occurrence restrictions on wh-words and foci, or special movement patterns which indicate that they form a natural class. For instance, in Italian, foci and wh-words cannot co-occur within the same clause (Rizzi, 1997), and multiple wh-questions are also not possible (Calabrese, 1984). A similar pattern is described for Quiegolani Zapotec (Black, 1994). By assuming that wh-words and foci form a natural class, restrictions of this type can be described straightforwardly: only one element of this class can be licensed within a clause. In Toba Batak, evidence that wh-words and foci form a natural class comes from movement. Though the language generally disallows multiple fronting—as in other Austronesian languages (Keenan and Comrie, 1977)—multiple foci or a focus and a wh-word can be simultaneously fronted (Erlewine, 2018).

A natural, and fairly common, hypothesis that has been advanced to account for the connection between wh-words and foci is to propose that they are both attracted by the same syntactic head. This claim has been made for Hungarian (É. Kiss, 1998a), Italian (Rizzi, 1997, 2001), and Mongolian (Onea and Guntsetseg, 2011), among other languages. Furthermore, some researchers have argued that both are attracted to this position because they bear the same syntactic feature. Throughout this paper, I refer to this idea as the Identity Hypothesis.

\[\text{(3)} \quad \textbf{The Identity Hypothesis:} \quad \text{The displacements of foci and wh-words are formally identical in the syntax. They move to the same syntactic position because their movement is driven by the same feature.}\]

Several versions of this hypothesis have been proposed in the literature. Much work has explicitly proposed that both wh-words and foci are attracted by a left-peripheral Focus head that searches for constituents that bear a \([\text{FOC}]\) (Bródy, 1990; Rizzi, 1997; Aboh, 2007; Aboh and Pfau, 2011; Aboh, 2016). Along the same line, Erlewine (2018) accounts
for the multiple fronting of foci and wh-words in Toba Batak by claiming that they are both “formally focused,” and thus share a feature which is subject to attraction by the same head. Finally, Horvath (1986) argues on the basis of Hungarian that [FOC] is a feature that is assigned to the constituent immediately preceding the verb, analogously to Case. She proposes that question wh-words must universally bear a focus feature at LF in order to be interpreted, and thus they move to the same syntactic position as other foci.

A challenge to this view, however, comes from the fact that wh-words move across more local quantificational foci in some languages. For instance, É. Kiss (1998b) shows that in Hungarian, though both exhaustive foci and wh-words must move to a preverbal position, only wh-words can appear in preverbal position when the two co-occur (4a). Moving an exhaustive foci and leaving the wh-word in situ is not grammatical (4b).

(4)  a. Kit látogatott meg csak Mari
      whom visited PREV only M.NOM
      ‘Who did only Mary visit?’

       b. *Csak Mari látogatott meg kit
          only M.NOM visited PREV whom
          Intended: Who did only Mary visit? 

   É. Kiss (1998b): 16-17

Though similar patterns have been observed for several unrelated languages, to my knowledge the consequences of this pattern for the theory of how foci and wh-words are represented in the grammar has not been explored.

In this paper, I extend this generalization beyond exhaustive foci using novel data from San Martín Peras Mixtec (henceforth, SMPM), an indigenous language of Southern Mexico. Though contrastive foci can be interpreted non-exhaustively in the language, the same pattern of movement holds: both foci and wh-words move to the same preverbal position in the language, but only wh-words can move there when they are in competition. Crucially, this pattern holds even when the wh-word is asymmetrically c-commanded by
the focus, and consequently is syntactically more distant from the probe (5b). Throughout this paper, I refer to this movement pattern as the “Wh-over-Focus Preference.”

(5)  

a. Ntsyá rí kitsí shín Marta __
    which CLF animal bought M.
    ‘Which animal did Marta buy?’

b. *Marta ŋá shín __ ntsyá rí kitsí
    M. she bought __ which CLF animal
    Intended: Which animal did Marta buy?

If we suppose that movement of both constituents is driven by the same feature, then there is an apparent violation of locality—assuming that probes must attract the closest active goal that they can Agree with (Chomsky, 2000), the most local goal (the focus) is being skipped by the probe in favor of a non-local goal (the wh-word).

In this paper, I argue for an alternative representation of wh-words and foci which accounts for this pattern, while maintaining a fundamental similarity between their respective movements. Specifically, I argue that focus movement and wh-movement are both the consequence of phrasal movement of a class of particles that are sensitive to focus alternatives (Cable, 2010; Branan and Erlewine, 2020, cf. Horvath, 2007). In addition to overt focus sensitive particles, SMPM has in its lexicon a null Question Particle (Q) (which c-commands wh-words) and a segmentally null focus sensitive particle (which c-commands contrastive foci and wh-congruence foci). Phrasal movement of any member of this class will trigger indirect movement of the alternative generating element in its c-command domain. In order to account for the Wh-over-Focus Preference, I propose that all focus sensitive particles in SMPM are lexically specified to bear a formal feature that is associated with their sensitivity to alternatives [ALT]. Q particles, however, represent a subclass of this larger class of focus sensitive particles. In addition to an [ALT] feature, they also bear the the feature [Q] (Cable, 2010).
Building on insights from the A-domain, I argue that these features are not independent of one another. Instead, I claim that they are arranged hierarchically with respect to one another in a feature geometry (cf. Abels, 2012; Foley and Toosarvandani, 2019). Specifically, I propose that the feature [Q] entails the feature [ALT]. This feature geometry allows us to simultaneously capture the fact that Q particles form a natural class with other focus sensitive particles, but also account for the fact that Q particles (and thus, wh-words) will move across more local focus sensitive particles (and thus, foci). This featural difference at the level of the focus sensitive particle builds on and older, common idea that wh-words themselves are a subtype of foci (Lee, 1999; Bošković, 2002; Sabel, 2000; Kim, 2006; Dong, 2009, a.o.).

While there is a fair amount of work on articulated probing in the A-domain (see e.g., Béjar, 2003; Béjar and Rezac, 2009; Oxford, 2019; Foley and Toosarvandani, to appear; Coon and Keine, to appear, a.o.), there is somewhat less work on articulated probing by probes that trigger phrasal movement in the Ā-domain (though see, Abels, 2012; Kotek, 2014; Hsu, 2017; Aravind, 2017). In this paper, I argue that the apparent non-local movement of wh-words in (5) can be accounted for using two tools from the A-domain: multiple searches within a locality domain (cf. Béjar and Rezac, 2009; Coon and Keine, to appear) and a constraint that economizes the valuation of probes (cf. Oxford, 2019; Coon and Bale, 2014; van Urk, 2015). I propose that a probe relativized to [u\textsc{alt}-uQ] will probe again past a focus particle in subject position, potentially finding a Q particle in object position, due to the fact that a focus particle alone does not completely satisfy its needs. The
object will value the probe if it constitutes a better match for the needs of the probe, and will subsequently be internally merged into the specifier of the probing head.

In addition to my proposal, I consider two alternative hypotheses to explain the Wh-over-Focus Preference. After providing some background on San Martín Peras Mixtec in §2, in §3 I consider whether wh-words and foci are moved by two distinct features. Under this hypothesis, wh-movement and focus movement are superficially similar, but are two distinct movements triggered by two distinct features, albeit to the same apparent position in some languages. Throughout this paper, I refer to this as the Disjoint Hypothesis.

(7) **The Disjoint Hypothesis**: Foci and wh-words are displaced by distinct syntactic features. The connection between them is epiphenomenal: both happen to move to similar surface positions, which creates the illusion that they are connected with each other.

While this alternative could account for some of the facts of SMPM, it encounters both empirical and typological problems. Empirically, I show that the Disjoint Hypothesis struggles to explain why focus movement is normally obligatory in Mixtec, except in the circumstances that a focus co-occurs with a wh-word. Typologically, I argue show that wh-words and foci very often surface in the same position, but the position to where they move varies across languages. Therefore, I argue that if we adopt the Disjoint Hypothesis, we are forced to resign ourselves to viewing the cross-linguistic syntactic connection between wh-words and foci as a mere coincidence. This, I argue, would be an unsatisfactory result.

After presenting my main proposal in §4, I return to the Identity Hypothesis in §5 and consider whether it could be maintained by assuming that the preference for wh-movement over focus movement is due to a Focus Intervention Effect (FIE) (Beck, 1996, 2006; Pesetsky, 2000; Kotek, 2019, a.o.). Given that an *in situ* wh-word in the scope of a
focus operator is ungrammatical in many languages, then it may be the case that fronting a focus instead of a wh-word is impossible in SMPM because it creates this marked structure. However, I show that, in SMPM, the Wh-over-Focus Preference holds even when the focus operator takes narrow, DP-level scope. Thus, I argue that the generalization that wh-words move across more local foci is not reducible to a FIE.

2 Mixtec and the Locality Problem

In this section, I provide the necessary background on focus in SMPM (§2.1), as well as its interaction with wh-movement (§2.2). Then, I identify a problem of locality if we assume the Identity Hypothesis: in SMPM, wh-words will always be attracted instead of foci, even if they are non-local (§2.3). Additionally, I argue that the properties of SMPM make it an ideal language to test the Identity Hypothesis (§2.4).

San Martín Peras Mixtec (ISO: JMX), also known as Tu’un Sávi or Tu’un Ndá’vi, is an Otomanguean language spoken by approximately 10,000 people in western Oaxaca, Mexico, near the state of Guerrero (Instituto Nacional de Estadística y Geografía, 2010). One of roughly 80 Mixtec varieties recognized by the Mexican government (Instituto Nacional de Lenguas Indígenas, 2008), it is classified as part of the Southern Baja dialect group (Josserand, 1983). All uncited data in this paper comes from my own fieldwork since 2017 with two speakers who live in California, as well elicitation with four speakers during a 2019 field visit to Ahuejutla, Mexico. All elicitation was conducted in Spanish.

The focus judgements in this paper were elicited in three primary ways. First, by establishing an explicit context and eliciting grammaticality judgements relative to that context. This was done either by asking for a translation from Spanish, or asking whether a Mixtec sentence would be acceptable in a particular context. Second, data on wh-congruence focus was gathered using printed images as a prompt to elicit responses to wh-questions.
that were posed in Mixtec. Third, I elicited semi-naturalistic speech by having speakers extemporaneously narrate stories, using storyboards designed to elicit particular focus constructions (Littell, 2010a,b; TFS Working Group, 2011; Grubic, 2014).

2.1 Focus in Mixtec

I assume, broadly speaking, that the focus of a sentence is the information that the speaker assumes the listener does not share (Jackendoff, 1972). Additionally, I assume that there are different types of focus which can have different properties, including foci that create a congruent answer to a wh-question, and foci that contribute an additional contrastive or corrective meaning (É. Kiss, 1998a). I assume that these types of focus are semantically distinguished by the scope of the focus operator (see §5.1 for discussion) (Rooth, 1992).

SMPM has a default VSO word order in out-of-the-blue contexts (Ostrove, 2018; Mendoza, 2020), as shown in (8).

(8) Kotô Agustina chichí
    likes A. avocado
    ‘Agustina likes avocados.’

While there has been little formal theoretical work on focus in Mixtec, it has been consistently observed that foci move to a preverbal position across a wide range of Mixtec languages (see, e.g. Alexander, 1988; DiCanio et al., 2018; Farris, 1992; Hedding, 2019; Hills, 1990; de Hollenbach, 2013; Johnson, 1988; Kuiper and Oram, 1991; Shields, 1988; Small, 1990; Zylstra, 1991). In SMPM, all types of foci are obligatorily fronted to a preverbal position. First, consider a case of wh-congruence focus, which is used to make a congruent answer to a wh-question (Rooth, 1992). Here, the word corresponding to the wh-word in the question must move to a preverbal position (9a). It is ungrammatical for it to remain in situ (9b).
(9) Context: What did the dog eat?
   a. Kôñù shishi rí __
      meat  ate  AML
      ‘It (an animal) ate the meat.’
   b. *Shishi rí  kôñù
      ate  AML meat
      Intended: It ate the meat.

It should be noted, in addition, that Broad Focus questions (which elicit all new information) often involve a fronted constituent in SMPM (10). This pattern of movement is well described for other languages, and has been variously referred to as “Partial Focus Movement” (Hartmann and Zimmermann, 2007), “Subpart of Focus Movement” (Fanselow and Lenertová, 2011), and “Anti-Pied-Piping” (Branan and Erlewine, 2020). Though a thorough investigation of this phenomenon in SMPM would take us too far afield, I refer the interested reader to Branan and Erlewine (2020) for a recent analysis of phenomena of these type.

(10) Context: Why is Maria upset?

   Nána  ŋá  nàkaba __
   mother her fell

   ‘Her mother fell down.’

   Contrastive foci, such as foci used to correct a previous misstatement, also move to a preverbal position (11a), and cannot be left in situ (11b).

(11) Context: I am eating in the house of my friend Gloria and she serves me some tortillas that she says her daughter Rosa made. Gloria’s husband tells her that she is wrong:

   a. U’un, Maria ŋá  tsyã __ shitã
      no  M. she made  tortilla
      ‘No, Maria made the tortillas.’

   b. *Maria  ŋá  tsyã __ shitã
      she  M. she made  tortilla
      ‘Maria made the tortilla.’
Contrastive foci are morphologically distinguished from non-contrastive foci in SMPM: they must be doubled by a clitic pronoun that agrees with their noun class. Simple congruence focus does not require clitic doubling (10a), however, congruence foci can be doubled by a pronoun if the context explicitly introduces alternatives, such as following an alternative question. Thus, the presence of a clitic pronoun seems to be related to the salience of alternatives. In this paper, I set aside the question of the precise syntax and semantics of doubled clitics as it is orthogonal to my main point. Instead, I focus on the fact that all foci undergo obligatory movement to a preverbal position.

Finally, several overt focus sensitive particles, corresponding to the English words even and only, also front to a preverbal position along with their associated focus in SMPM. These particles, like their English equivalents, require some focused element in their c-command domain.

(12)  

\[
\begin{align*}
\text{a. } & \text{Inta } \text{ñá maestra } \text{ñá ká'an } \text{tu'un } \text{sá'á} \text{ only CLF teacher she speaks language Spanish} \\
& \text{‘Only the teacher speaks Spanish.’}
\end{align*}
\]

\[
\begin{align*}
\text{b. } & \text{*Ká'an inta } \text{ñá maestra } \text{ñá tu'un } \text{sá'á} \\
& \text{speaks only CLF teacher she language Spanish} \\
& \text{Intended: Only the teacher speaks Spanish.}
\end{align*}
\]

Several overt focus sensitive particles can also attach to VPs.

(13)  

\textit{Context: Pedro’s brother is very lazy.}

\[
\begin{align*}
\text{Niná kíshi } \text{ñañi } & \text{Pedro} \\
& \text{only sleeps brother P.} \\
& \text{‘Pedro’s brother only sleeps.’}
\end{align*}
\]

Given that SMPM is a predicate initial language, it is not immediately clear if particles of
this type trigger movement. Some evidence that they do not is that when a VP is focused with a PP adjunct, the PP adjunct remains in a position after the subject.

(14) Context: How does Mariana stay in shape? Does she run?

U’un, niná ŋá ishutsya ŋá ini tskwï
no only she swims she in water
‘No, she only swims in the ocean.’

While focus sensitive particles generally undergo movement in the language, there is one particle that I am aware of—và—which can attach to foci but does not trigger overt movement.

(15) Context: The dog ate some tortillas, right?

Ũ’ũ, kônù ŋá shàshi ri __ ra shàshi ti ri chichí và
no meat NEUT ate AML and ate also it AML avocado FOC
‘No, it ate the meat, and it also ate an avocado.’

While this particle can sometimes be used to mark new information (as in 15), it does not seem to be a preferred strategy to mark new information among the speakers that I have consulted. Indeed, according to Cisneros (2020), the cognate particle in closely related Cuevas Mixtec has several distinct uses: it can mark new information, it is used in the formation of free choice indefinites, it can denote an ongoing event when attached to a verb, and can be used to signal that an utterance runs counter to the speakers expectations. As this particle does not appear to trigger movement, I set it aside for the rest of this paper, and leave to future work a more complete analysis of its distribution and how it interacts with focus. I refer the interested reader to Cisneros (2020) §5.3 for more details on the use of this particle in Cuevas Mixtec.

There are three diagnostics that demonstrate that focus displacement in SMPM is a type of syntactic movement. First, foci cannot be moved out of islands, such as subjects
Second, parts of idiomatic expressions can be focused and retain their idiomatic meaning. This can be seen, for example, with the expression *shashi yú’ū* which literally means “to eat mouth,” but which has the idiomatic meaning of “to kiss.” When part of the idiom is focused, it can still be interpreted idiomatically.

19) Context: I go to a dinner party with my son Juan. At the party, Juan meets up with his girlfriend Maria. When we eat dinner, I can’t find Juan. When I go looking for him after dinner, I find him and Maria kissing. Later, at home, my wife notices that Juan looks hungry. She asks: “What did Juan eat at the party?”

*Yú’ū*  *Maria*  shàshi rà
mouth M.  ate he
‘He kissed Maria.’ (literally, He ate *Maria’s mouth*)
Finally, reciprocals can be bound in a fronted focus position (20). Evidence that they must be bound by the subject comes from the fact that making the subject singular leads to ungrammaticality (21).

(20) Context: Did you both give a photo to your mother?

U’un, ndà’ǎ tá’an míí ndù tâshin ndù ŋà no to each. other EMPH IPLEXCL gave IPLEXCL it
‘No, we gave a photo to each other.

(21) *U’un, ndà’ǎ tá’an míí ndù tâshin Juan ŋà no to each. other EMPH IPLEXCL gave J. it
Intended: No, Juan gave a photo to each other.

Taken together, these diagnostics provide strong evidence that focus displacement is a syntactic phenomenon in SMPM, as has been argued for other languages.

Additional evidence that focus movement is syntactic comes from clauses with multiple foci. As expected, assuming standard syntactic locality applies, only the structurally highest focus can front to a preverbal position.

(22) Context: One of my friends doesn’t really like mushrooms—he almost never eats them. However, he really likes one mushroom in particular: the deer horn mushroom. I relate this to my friend Benjamín and he says: “How surprising! Which of your friends eats only deer horn mushrooms?”

a. Juan rà shishi __ inta shi’ǐ ntsiki usu
   J. he eats only mushroom horn deer 'Juan eats only deer horn mushrooms.

b. *Inta shi’ǐ ntsiki usu shîshi Juan rà __
   only mushroom horn deer eats J. he
   Intended: Juan eats only deer horn mushrooms.

I take this to be additional support for the claim that focus movement is syntactic and thus
is sensitive to syntactic constraints on movement such as locality.²

2.2 Interactions with Wh-Movement

Wh-words move obligatorily to a preverbal position in SMPM, as in other Mixtec languages (e.g. Caponigro et al., 2013).

(23) a. Yóó shástì _ kwi’ì?
    who ate fruit
    ‘Who ate the fruit?’

b. Nã shì Martà __
    what bought M.
    ‘What did Martà buy?’

c. Nashá yávi yá __
    how.much costs it,NEUT
    ‘How much does it cost?’

d. Ntsyá chíchi ndó __
    where bathe 2pl.

²An anonymous reviewer points out that in Hungarian—another language claimed to move foci syntactically—either focus can move in clauses with multiple foci (É. Kiss, 2002). These distinct movement possibilities determine which focus takes highest scope. I assume, following (É. Kiss, 2002, §3.4), that the Hungarian VP is “flat” and that neither the subject nor object asymmetrically c-commands the other. É. Kiss (2002) argues in favor of this hypothesis on the basis of word order, binding, and the lack of superiority violations in multiple wh-questions. If correct, then this hypothesis would correctly predict that either the subject of object could be focus fronted, given that neither asymmetrically c-commands the other and thus neither acts as an intervener.

The same flat structure structure cannot be posited for SMPM, given evidence from binding which suggests that the subject asymmetrically c-commands the object. For instance, an R-expression in object position cannot be bound by a co-referential subject pronoun (ia). However, an R-expression embedded in a possessive DP can co-refer to an object pronoun. If the subject and object mutually c-commanded one another, we would expect co-reference in (ib) to also be ungrammatical.

(i) a. Kôni ŋâ; se’e Maria_s/j
    loves she child Maria
    ‘She, loves Maria_s/j’s child.’

b. Kôni Maria_s se’e ŋâ_s/j
    loves she child Maria
    ‘Maria, loves her_s/j child.’

Therefore, I assume that the difference in the acceptability of fronting object foci across subject foci in the two languages is reducible to the different structural relationship between subjects and objects in the two languages.
‘Where do you (pl) bathe?’

Multiple wh-questions are ungrammatical in the language, as has been claimed for other varieties of Mixtec.³

(24)  Context: I have three children. I give them each 5 dollars to buy some candy in the store. Later on, I see three candy wrappers in the trash. I ask:

a. *Yóó ndó shǐn  ntsyâ  dúlse?
   who 2PL bought which candy
   Intended: Which of you bought which candy?

b. *Yóó ndó ntsyâ  dúlse  shǐn?
   who 2PL which candy bought
   Intended: Which of you bought which candy?

Wh-words and foci move to the same syntactic position in SMPM, as has been proposed for other languages (Rizzi, 1997). This can be shown based on the fact that they appear in identical positions with respect to several syntactic landmarks. First, they can both surface above fronted negative indefinites (25) and preverbal manner adverbs (26).

(25)  a. Yóó ko  ñá’a  niki’i
   who  NEG thing bought
   ‘Who bought nothing?’

b. U’un Maria  ñá  ko  ñá’a  niki’i
   no  M.  she  NEG thing bought
   ‘No, Maria bought nothing.’

(26)  a. Yóó ni’i  shǐnu
   who strong runs
   ‘Who runs fast?’

b. Teresa  ñá  ni’i  shǐnu
   T.  she strong runs

³Four of the five speakers that I have consulted about multiple wh-questions have consistently rejected them. One speaker, however, fairly consistently accepts them. At this time, I do not know what accounts for this inter-speaker variation.
Second, temporal adverbs—which can normally appear either preverbally or sentence-finally—must occur sentence finally when combined with a fronted wh-word (27) or focus (28).

(27)  
   a. Yóó nishi’i koni who died yesterday ‘Who died yesterday?’  
   b. *Koni yóó nishi’i yesterday who died Intended: Who died yesterday?  
   c. *Yóó koni nishi’i who yesterday died Intended: Who died yesterday?

(28)  
   Context: Vitorino broke his hand yesterday.  
   a. U’un, Juan rà nitá’no ndâ’á koni no J. he broke hand yesterday ‘No, Juan broke his hand yesterday.’  
   b. ??U’un, koni Juan rà nitá’no ndâ’á no yesterday J. he broke hand Intended: No, Juan broke his hand yesterday.  
   c. *U’un, Juan rà koni nitá’no ndâ’á no J. he yesterday broke hand Intended: No, Juan broke his hand yesterday.

   Moreover, fronted wh-words and foci are in complementary distribution with one another. Movement of a focus to a position between a wh-word and the verb is completely ungrammatical (29a). Instead, the most natural way to form this question is to leave the focus in situ (29b).

(29)  
   Context: My friend Benjamín and I went a party where everyone brought some
food or drink to share. We know who brought the meat but we are curious who
brought the tortillas. When we go to ask the host, Benjamín mistakenly asks him,
“Who brought the meat?” I turn to Benjamin and say, No:

a. *Yóó shitā yá shin’ī __ __
   who tortilla NEUT brought
   Intended: Who brought the tortillas?

b. Yóó shin’ī __ shitā yá
   who brought tortilla NEUT
   ‘Who brought the tortillas?’

Crucially for the upcoming discussion, a focus will stay in situ regardless of whether it is
the subject or the object.

(30) Context: My friend Benjamín and I went a party where everyone brought some
food or drink to share. We know what Maria brought, but we didn’t see what
Marta brought. When we go to ask the host, Benjamín mistakenly asks him,
“What did Maria bring?” I turn to Benjamín and say, No:

   Nǎ kishashi Marta ñá __
   what brought M. she
   ‘What did Marta bring?’

For the sake of completeness, it is also necessary to rule out the possibility of a focus
fronting to a position above a wh-word. At first glance, this seems to be possible, how-
ever, there are three reasons to show that in examples like (31), the object has not under-
gone syntactic focus movement, but is instead acting as a contrastive topic.

(31) Context: Same as (29)

   Shitā yá, yóó shin’ī
tortilla NEUT who brought
   ‘The tortillas, who brought them?’

First, we have seen above that syntactic focus fronting is obligatory in SMPM, except
in cases where it is blocked by a wh-word. However, fronting a constituent to a position above a wh-word is always optional, as can be seen by comparing (31) to (29b). This suggests that this is a distinct syntactic phenomenon.

Second, if the pre-wh-word constituent were a focus attracted by the same head that attracts wh-words, then we would expect their relative order to depend on which is the subject and which is the object. In fact, contrastive topics always appear above wh-words, regardless of their base position.

(32) Same as (30)

Marta ñá ná kishashi
M. she what brought
‘Marta, what did she bring?’

Third, and perhaps most importantly, certain expressions, such as negative indefinites (33) and universally quantified DPs (34) are restricted from appearing in this pre-wh-position.

(33) Context: I had 10 guests at a potluck at my house but only 9 dishes. Looking for the person who didn’t contribute, my friend Benjamin begins to ask dish by dish, “Who brought the meat? Who brought the salsa...” I impatiently ask:

*??Ko-ñá’a ñá yóó ní-shini’i
NEG-thing NEUT who NEG-brought
Intended: Nothing, who brought?

(34) As my friend and I walk past a school, we see that all of the children are leaving one of the classes crying. It seems as though one of the teachers was in a bad mood and scolded them. We are worried, so we go to ask the principal what happened. My friend asks: ‘Which teacher scolded the children?’ I interject:

*Ntsi’i na bali ntsyå rà maestro nikàná’a shi’ìn
all they small which he teacher scolded with
Intended: All the children, which teacher scolded?

These restrictions would be surprising if movement to a position before the wh-word was an optional type of focus movement, as *no one* and universally quantified DPs can both undergo focus fronting with no restriction.

(35) Context: I had a potluck, but my friend Maria didn’t bring anything to share. My friend Benjamin notices that I am frustrated about something and tries to cheer me up, saying: “Well, at least Maria brought her delicious salsa.” I respond, No:

\[
\text{Ko-ñà’a ñà ní-shini’í Maria} \\
\text{NEG-thing NEUT NEG-brought M.} \\
\text{‘Maria brought nothing.’}
\]

(36) Who did they scold at the school today?

a. \text{Ntsi’i nà báli nikàná’a nà shí’in} \\
\text{all they small scolded they with} \\
\text{‘They scolded all the children.’}

The same contrast holds in Italian (Rizzi, 1997). While negative indefinites and universally quantified DPs cannot be topicalized in Clitic Left Dislocation structures, they can be focused. Collectively, these facts suggest that the position above wh-words is reserved for contrastive topics, which must be referential. Thus, I assume the left-peripheral structure for SMPM shown in (37). Both foci and wh-words move to a preverbal position, which I assume to be the specifier of a null C. In addition to this position, there is at least one higher peripheral position that is reserved for topics. See Macaulay (1996) and Aissen (1992) for similar proposals about Chalcatongo Mixtec and Mayan, respectively.
2.3 The Locality Problem

Now that we have shown that foci and wh-words move to the same syntactic position in SMPM—as in other languages—the next step is to propose an account of how they are represented in the syntax in order to derive this pattern. Let us begin with the the simplest view on the relationship between wh-words and foci that has been proposed in the literature: wh-words and foci are both marked with a feature, $[\text{FOC}]$, which triggers their movement (Horvath, 1986; Aboh, 2016). According to this Identity Hypothesis, wh-words and foci are formally identical in the syntax and the motivation for focus movement and wh-movement is the same.

This view makes a concrete prediction about movement that will be fruitful to explore. Specifically, assuming an Attraction-based theory of movement, it predicts that in languages where there is one designated position for foci and wh-words, and only one thing can move to that position, the structurally highest constituent marked $[\text{FOC}]$ should move—a type of superiority effect. If wh-words and foci are attracted by the same probe, then we expect that probe to attract the most local goal within its domain (Chomsky, 1993, 2000, 2001; Pesetsky, 2000; Rizzi, 1990). A probe should not be able to skip any potential goals in order to find a non-local goal. This explains, for instance, why fronting an object wh-word across a more local subject wh-word is generally ungrammatical in English (Kuno and Robinson, 1972; Chomsky, 1973).

Consequently, under the view that wh-words and foci are attracted by the same fea-
ture, we expect to find languages where wh-words and foci display the same pattern. That is, we expect some hypothetical language to move wh-words when they are in subject position and a focus is in object position (38a), but to move the focus when it is structurally higher than a wh-word (38b).

(38)  
   a. Subject Wh-word Moves, Object Focus Remains in situ
       \[WH \ldots \_\_ FOC\]
   b. Subject Focus Moves, Object Wh-word Remains in situ
       \[FOC \ldots \_\_ WH\]

The Identity Hypothesis makes another, related prediction: there should be no structurally-insensitive preference to front either wh-words or foci. We don’t expect, for example, wh-words to be consistently attracted, even if they are structurally lower than a focus in the same clause (39).

(39)  
   Object Wh-Word Moves, Subject Focus Remains in situ
   \[*WH \ldots \_\_ FOC \_\_\]

Let’s scrutinize these predictions. As we have seen, SMPM is a language with obligatory wh-movement and obligatory fronting of foci to a position before the verb. Furthermore, we have reviewed data that suggest they move to the same syntactic position. However, there is a puzzle that arises when we consider clauses with more than one of these elements: wh-movement always takes precedence over focus movement, even if the focus would be a more local goal to the probe, as in (40).

(40)  
   Ntsyấ rí kitsî́ shîn \textbf{Marta} \_
       which CLF animal bought M.
   ‘Which animal did \textbf{Marta} buy?’
Thus, there is a “Locality Problem” that arises when we consider how wh-words and foci interact in SMPM: namely, attraction to the preverbal position does not seem to respect locality constraints. Instead, there seems to be a general preference to move wh-words.

As mentioned in section 1, this locality problem is not restricted to SMPM. For instance, Hungarian is a language where both wh-words (41a) and contrastive foci (41b) move to an immediately preverbal position. Like in SMPM, there is only one syntactic position into which foci and wh-words can move, leading many to argue that focus-fronting and wh-movement should be understood as a unified phenomenon (Horvath, 1986).

(41)  

a. Hol jártál a nyáron?
   where went.you the summer.in
   ‘Where did you go in the summer?’

b. **Olaszországban jártam**
   Italy.to went.I
   ‘It was **Italy** where I went.’

É. Kiss (1998a): 249-250

As there is only one preverbal focus position in Hungarian, the Identity Hypothesis predicts that the highest element marked \([\textit{FOC}]\) will move there, be it a wh-word or a contrastive focus. This prediction is not borne out, however. In sentences with both a wh-word and a focus, only wh-words can move to the preverbal position, regardless of whether they originate as the subject (42) or the object (43). Foci will never move instead of a wh-word, even if they originate in a structurally higher position.

(42)  

a. Ki látogatta meg csak **Marit**
   who visited \(\text{PREV} \text{only } \text{M.ACC}\)
   ‘Who visited only Mary?’

b. *Csak **Marit** látogatta meg ki
   only \(\text{M.ACC} \text{visited } \text{PREV}\) who
   Intended: Who visited only Mary?
Thus, like SMPM, Hungarian seems to display a general preference for wh-movement over focus fronting. Generalizing across both languages, we can state the hypothesis in (44).

(44) **The Wh-over-Focus Preference**: In languages where wh-words and foci are in competition for a single position, only wh-words can front to that position when they co-occur.

This generalization also seems to extend beyond SMPM and Hungarian. A similar preference to move wh-words over foci has been reported in Basque (Hualde and Ortiz de Urbina, 2003, pg. 495) and Georgian (Borise and Polinsky, 2018, pg. 5). While suggestive of a broader cross-linguistic preference to move wh-words over foci, both these cases require caveats. In Basque, some speakers reportedly find wh-words and foci co-occurring within the same clause unacceptable (Ortiz de Urbina, 1999, pg. 315). As for Georgian, Borise and Polinsky (2018) argue that wh-words and foci occupy a preverbal position for prosodic, rather than syntactic reasons. I hope that future investigation into these languages and others clarifies the extent to which the Wh-over-Focus Preference holds cross-linguistically.

Thus, we find an asymmetry that is unexpected if we adopt the Identity Hypothesis: in cases where both foci and wh-words could in theory move, wh-words take precedence. This strongly suggests that there is something about wh-words that makes them special, and that makes them more likely to move than foci. I have not yet seen an at-
tested language that moves either foci or wh-words, depending on which is structurally highest (38). I take this as evidence that the Identity Hypothesis cannot be maintained for SMPM, nor for other languages which display this preference.

2.4 The Contribution of Mixtec

While, on its face, Hungarian seems to be an adequate language to scrutinize the predictions of the Identity Hypothesis, in fact, an idiosyncrasy of the language introduces an important confound. Hungarian only moves foci that are interpreted exhaustively; non-exhaustive foci remain in situ (É. Kiss, 1998a; Horvath, 2007). Thus, it is not clear whether the language displays focus-movement, per se, or whether exhaustive identification is the crucial driver of movement, as proposed in Horvath (2007). Further complicating the picture, fronted wh-words do not have to be interpreted exhaustively in the language, as evidenced by the the felicity of “mention-some” questions (Cable, 2008).

(45) a. Hol vehetek újságot itt a környéken?
   where.I.can.buy newspaper.ACC here the vicinity.on
   ‘Where can I buy a newspaper around here?’

b. Melyik számok páratlanok?
   which numbers odd.PL
   ‘Which numbers are odd?’

Cable (2008): 10

4One language that may display this pattern is Amahuaca, a Panoan language spoken in Peru and Brazil (Clem, 2019). Clem reports that is possible for transitive subjects to move to a position in front of an interrogative particle (analyzed as the head of C), leaving a wh-word in object position.

(i) Jan hinan=ra tzova vuchi=hax
   3SG.GEN dog.ERG=INT who find=TAM
   ‘Who did his dog find?’

(Emily Clem, p.c.)

Clem notes that the context in which (i) appears plausibly allows the subject to be interpreted as a focus, possibly instantiating the prediction of the Identity Hypothesis. However, she cautions that she has not specifically investigated the interaction between wh-movement and focus movement in the language.

If further investigation of Amahuaca indicates that movement of foci and wh-words depends on the structural position of each, then it might be an example of a language where they are indeed represented identically. Alternatively, anticipating the proposal in section 4, it might be the case that the probe on C in Amahuaca is only relativized to [ALT], thus making foci and wh-words an equally good match for the needs of the probe.
This discrepancy between the interpretation of fronted foci and fronted wh-word leads (Cable, 2008) to suggest that they cannot be the same phenomenon. Though, as Cable himself recognizes, arguing that wh-movement and focus-fronting are distinct “re-problematizes the structural similarities between wh-questions and focus constructions” (pg. 12). By only looking at Hungarian, we have an incomplete picture of the cross-linguistic facts, and don’t have the ability to conduct the crucial experiment to test the Identity Hypothesis without introducing a critical confound. Of crucial importance, then, will be demonstrating that SMPM does not introduce the same confound, and for that reason makes it an ideal language to investigate the relationship between focus (broadly construed) and wh-movement.

All foci move in SMPM, not just foci that are interpreted exhaustively. This can be shown in three distinct ways. First, there is no contrast in the movement behavior of focus sensitive particles which trigger an exhaustive interpretation (46), and those that are incompatible with an exhaustive interpretation (47).⁵

(46) Context: Does Margarita eat meat?
   a. U’un, inta yibá shishi ñá __
      no only vegetables eats she
      ‘No, she only eats vegetables.’
   b. *U’un, shishi ñá inta yibá
      no eats she only vegetables
      Intended: No, she only eats vegetables.

(47) Context: We are rushing to make enough tortillas before a party. Pedro, who famously hates cooking, is helping.

---

⁵One of the two speakers that I consulted about these examples occasionally accepted leaving ntsy a (even) and its associate in situ. However, these judgements were variable across elicitation sessions, leading me to conclude that it is at best marginally acceptable. Moreover, both speakers consistently accepted moving ntsy a, suggesting that it can be attracted, in contrast to Hungarian. I leave exploration of the precise reason why leaving it in situ is sometimes accepted as grammatical by some speakers to future work.
A contrast between these two particles is present in Hungarian—*only* and its associate must undergo fronting in Hungarian, but *even* and its associate cannot—further indicating that only certain types of foci undergo movement (see Horvath, 2007, pg. 121-122).

Second, both congruence and contrastive foci can felicitously be followed up with a continuation which adds another element to the focused set (48a). Crucially, as excepted, adding a focus sensitive particle that explicitly triggers an exhaustive interpretation makes the continuation infelicitous (48b).

(48)  
*Context: Teresa chatted with Esteban at the school yesterday.*

a. U’un, shi’i **Margarito** nda’a tu’un ŋá. Sá ti shi’i **Juan.**
   No with M. hand word she so also with J.
   ‘No, she chatted with **Margarito.** Also, with **Juan.**

b. U’un, inta shi’i **Margarito** nda’a tu’un ŋá. #Sá ti shi’i **Juan.**
   No only with M. hand word she #so also with J.
   ‘No, she only chatted with **Margarito.** #Also, with **Juan.**

Third, truth value judgements confirm that a fronted focus can serve as a partial answer (49a). Once again, this contrasts with sentences with an overt focus sensitive particle that triggers an exhaustive interpretation, such as *only* (49b).

(49)  
*Context: I went to a party with Francisca. There was a lot of food there. She ate some mole, beans, tortillas, and some cake for dessert. If someone asks me: ‘What did Francisca eat ate the party?’ would the following answers be true?:*

a. TRUE: **Ndayají vá’a** shishi ŋá
   broth good ate she
‘She ate mole.’

b. FALSE: Inta ndayají vá’a shishi ñá
only broth good ate she
‘She only ate mole.’

Thus, given that all types of foci undergo syntactic movement in SMPM, and there are no apparent semantic restrictions on the foci that front, it provides an ideal test case to investigate the predictions of the Identity Hypothesis. In the next section, I explore an alternative approach to the relationship between wh-words and foci: they have completely non-overlapping features and are attracted by distinct probes.

3 Disjoint Features

In the previous section, I outlined a generalization about the interaction between wh-words and foci: in several unrelated languages, wh-movement takes precedent over focus movement when the two are in competition for a single position. There is, however, another hypothesis that should be considered: wh-words and foci are not attracted by the same head, but are instead attracted by two distinct heads which are in complementary distribution with one another. This analysis could feasibly provide an explanation for the preference of wh-movement over focus movement in SMPM, if we assume that the Wh-over-Focus Preference actually reflects a preference to select a head that triggers wh-movement. This hypothesis is especially important to consider given the rich literature which argues for an articulated left-periphery, where the domain of complementizers does not consist of a single head C, but rather a series of functional projections—among them TopicP, FocusP, and InterrogativeP—each of which has a designated function (e.g. Rizzi, 1997, 2001; Aboh, 2016; Frascarelli and Puglielli, 2007a,b; Shlonsky and Bocci, 2019). Broadly speaking, the consensus within the cartographic literature has been that
wh-words and foci are attracted by the same head, given the strong correlation between them and the fact that they are in complementary distribution (Rizzi, 1997; Aboh, 2016; Frascarelli and Puglielli, 2007a). However, the logic of the Cartographic Program allows us to consider the possibility that SMPM has multiple attracting heads within the C domain, one which moves wh-words and one which moves foci.

In this section, I explore this alternative hypothesis as a means to explain the complementarity of wh-words and foci, as well as the general preference to move wh-words within SMPM. This hypothesis—which I refer to as the Disjoint Hypothesis—accounts for the locality problem outlined in the previous section by proposing that wh-words and foci do not share any features, nor are they attracted by the same head, despite the fact that they appear to surface in similar syntactic positions. Instead, this hypothesis points to differing licensing conditions as the source of their differing behaviors. As I will show, this alternative can account for some of the facts of SMPM, but it encounters two problems, one empirical and one typological. Empirically, this analysis struggles to explain the fact that focus movement is obligatory in SMPM when no wh-word is present. Typologically, the analysis offers no principled explanation for the robust syntactic correlation between wh-words and foci. Such an account would be forced to conclude that their connection is epiphenomenal—that is, both wh-words and foci are independently attracted by two distinct probes on C, creating the illusion that they are connected with one another. Consequently, this analysis would force us to abandon a deeper generalization about the connection between wh-words and foci, and thus should be dispreferred to an analysis that offers a principled explanation for the connection between the two.
3.1 Two Heads in the C Domain

Under the Disjoint Hypothesis, SMPM has two distinct probe-bearing heads in the C domain which are in complementary distribution with one another. One of these probes attracts foci and one attracts wh-words. Under this analysis, foci bear a feature \([\text{FOC}]\) which can enter into an agreement relationship with a focus probe, while wh-words bear a distinct feature \([\text{WH}]\) which can enter into agreement with a wh-probe. These two distinct probes, along with the disjoint features of foci and wh-words, ensure that there is no single probe that is searching for both foci and wh-words, and thus neither will act as an intervenor for the other.

Under this analysis, the apparently non-local movement of wh-words is attributable to the fact that the probe which attracts wh-words does not interact in any way with foci. Anything that is not a wh-word will not intervene due to the relativization of the probe (50a). The fact that a wh-word moves instead of a more local focus is unsurprising—the focus simply doesn’t have the feature that the probe is looking for. Under this analysis, focus movement is driven by a different probe, which is relativized to \([\text{uFOC}]\). When the focus probe is merged, it will move the most local constituent marked as \([\text{FOC}]\), ignoring any constituent that doesn’t bear a focus feature (50b).

(50)  

\[(50) \quad \begin{array}{ll}
\text{a.} & \text{WhP} \quad [u\text{WH}] \quad \text{v} \quad \text{DP} \quad [\text{FOC}] \quad \text{DP} \quad [\text{WH}] \\
\text{b.} & \text{FocP} \quad [u\text{FOC}] \quad \text{v} \quad \text{DP} \quad [\text{WH}] \quad \text{DP} \quad [\text{FOC}]
\end{array}\]

In order to account for the complementarity of wh-words and foci in the left-periphery,
this hypothesis would have to stipulate that the probes which target foci and wh-words, respectively, are in complementary distribution with one another. Assuming that either of the two probes could be merged in any given derivation, then there must be some mechanism to ensure that the wh-probe is always merged when a wh-word is present, forcing it to be attracted instead of the focus. Put differently, the derivation needs to be ruled out where the focus probe is merged, attracting a focus in subject position and leaving an object wh-word \textit{in situ}, a configuration which is ungrammatical in SMPM (51).

(51)

One reasonable way to rule out this configuration would be to say that wh-words have a licensing requirement which forces them to enter into an agreement relationship with a wh-probe (cf. the \textit{wh-criterion} May, 1985; Rizzi, 1996). The effect of this licensing requirement is that a wh-probe must be merged in any derivation that has a wh-word, or else the wh-word will not be properly licensed. Merging the focus probe instead will cause the derivation to crash, as the wh-word will not have its licensing needs met (cf. the \textit{Greed Principle}, Chomsky, 1995; Bošković, 1995). Thus, wh-words will be attracted when they are present in a derivation, while foci will either be attracted or left \textit{in situ}, depending on which probe is merged.

A crucial assumption of this analysis would be that foci do not have the same licensing needs as wh-words in SMPM. That is, foci can surface in a derivation with a wh-probe (as in 50 above), but wh-words cannot appear in a derivation with a focus probe, as
they will remain unlicensed (51). Thus, on this account, because wh-words have stricter licensing requirements than foci, they will be attracted in derivations where both occur, accounting for the the locality problem outlined in the previous section.

This assumption encounters an empirical problem, however, when we recall that focus movement is obligatory in SMPM (§2.1). If the focus probe and the wh-probe are in free variation, and only wh-words are subject to a licensing requirement, then we expect focus fronting to be optional—foci will move if and only if the focus probe is merged (52a) and will remain in situ if the wh-probe is merged (52b). Crucially, the assumption required to explain the preference for wh-words—the difference in licensing requirements between wh-words and foci—forces us to conclude that there should be no problem if a focus does not enter into an Agreement relationship with any probe (52b). It is, however, always ungrammatical to leave a focus in situ if there is no wh-word in the clause.

(52)  a.  

Thus, assuming disjoint features confronts an empirical problem: how can we ensure that focus movement will be obligatory if there is no wh-word present in the derivation, but prohibited if there is one? The Disjoint Hypothesis can explain the preference for movement of wh-words, but the necessary assumptions for that analysis lead us to predict that focus movement should be optional in cases with no wh-words, contrary to fact.
3.2 Missing a Deeper Generalization

Recall the robust cross-linguistic generalization that we began with: in languages where wh-words and foci both move, they seem to move to the same position. Thus, a desideratum of any analysis should be to explain this correlation in a principled way. The alternative analysis sketched in the previous subsection supposes that the reason that wh-words and foci move to the same position is because there happen to be two attracting heads in the C domain in SMPM: one that targets foci and one that targets wh-words. That is, the connection between these two elements boils down to the particular lexical items and probes that are available in the language.

However, when we consider a wider variety of languages, we see that although wh-words and foci consistently move to the same position, the position within the clause that they move to varies from language to language. In some languages, foci and wh-words both move to the complementizer domain (e.g. Italian & Gungbe, Rizzi, 1997; Aboh, 2007) and in other languages, the move to a position within the inflectional domain immediately before the verb (e.g. Hungarian & Malayalam, É. Kiss, 2002; Jayaseelan, 2001). Each of these languages instantiates the correlation between wh-movement and focus movement, yet there is no consistent position to which they move. This is important, because it indicates that the connection between wh-words and focus is not epiphenomenal. That is, it is not the case that they are both connected to a particular position (say, spec-CP), and that these two independent connections create the illusion that they are correlated with one another. Rather, wh-words and foci seem to be tightly connected to one another, regardless of the structural position that they ultimately move to.

Given the fact that wh-words and foci can, in principle, appear in multiple positions throughout the clause—especially in the C domain and v domain—if we adopt the Disjoint Hypothesis then we expect to find languages where the two probes that attract them
are hosted on distinct heads, even in distinct parts of the clause. In actuality, the general-
ization that wh-words and foci move to the same position is quite cross-linguistically robust. It is unlikely the case that it is merely a coincidence that in so many languages the probe searching for wh-words and the probe searching for foci happen to be merged in the same position. Adopting any analysis that cannot explain that larger generalization is, it seems to me, theoretically unsatisfying.

Given this desideratum—to explain the syntactic connection between wh-words and foci in a principled way—in the following section, I propose an analysis arguing that movements of both are triggered by the same head. As I will show, this analysis attributes the fact that wh-words and foci seem to move to the same position by supposing that they are attracted by the same head, which can surface in different positions throughout the clause in different languages. Furthermore, I claim that the features associated with each can account for the fact that wh-words move instead of more local foci.

4 Movement of Focus Sensitive Operators

In the previous two sections, I have outlined problems with two possible hypotheses on the relationship between wh-words and foci. Positing an identical featural representation gives rise to incorrect predictions when we consider that in several languages—SMPM among them—wh-words will move across more local foci. Positing two disjoint heads that are in complementary distribution faces both empirical problems and conceptual problems. On the one hand, it can’t straightforwardly explain why focus movement should be obligatory absent a wh-word, and on the other hand, it provides no principled explanation for why wh-words and foci appear to move to the same position in language after language.

In this section, I advance a general proposal for movement of focus sensitive parti-
cles. This proposal, along with certain assumptions about how probing proceeds in the Á-domain, accounts for the fact that wh-words and foci move to the same syntactic position, but there is a general preference to move wh-words across more local foci. A key component of my proposal relies on the notion that both foci and wh-words introduce focus alternatives, and both can appear in the scope of particles that are sensitive to those alternatives (Rooth, 1992; Kratzer and Shimoyama, 2002; Ramchand, 1997; Cable, 2010; Kotek and Erlewine, 2016, a.o.). I argue that movements of both foci and wh-words are indirect, triggered by attraction of focus sensitive particles that c-command them. This class of focus sensitive particles undergo syntactic movement in some languages, simultaneously fronting the alternative generating constituent that they c-command.

I propose that a single articulated probe on C attracts focus sensitive particles in SMPM, which I argue bear a formal syntactic feature $[\text{ALT}]$ (indicating that they are sensitive to alternatives). When these particles move, the focus or wh-word that they c-command moves along with them. However, the focus sensitive particle that must c-command wh-words (Q), bears an additional formal syntactic feature $[\text{Q}]$. As the probe is relativized to look for both $[\text{ALT}]$ and $[\text{Q}]$, it can interact with multiple goals in some circumstances. Building on some work on probing in the A-domain, I argue that there is a preference for the probe to be valued by the goal which bears features which most closely match the probe’s needs (cf. Oxford, 2014, 2019; Coon and Bale, 2014; van Urk, 2015). Thus in cases where both a Q particle and another focus sensitive particle are candidates for movement, the Q particle will be attracted instead of the other particle. This proposal is based on the idea that syntactic Agreement can be separated into two component parts: a search process and a valuation process (Chomsky, 2000). While the search process is local, the probe does not have to be valued by the most local goal, and thus the subsequent Internal Merge operation does not have to be local.
4.1 QP Movement

Recent work by Cable (2010) proposes that the phenomenon traditionally called ‘wh-movement’ is in fact movement of a Question Phrase (QP) that contains a wh-word. When the phrase headed by a Question Particle (Q) is attracted, the wh-word will also move, as it is part of the larger QP constituent (53). According to Cable’s proposal, wh-words must appear in the scope of a Q particle due to the fact that they have a particular semantic deficiency: while they introduce focus alternatives, they do not denote anything (Beck, 2006). Thus, according to this view, wh-words like who do not have a denotation, but instead simply introduce a set of alternatives which are presupposed to be human: {Juan, Eraclio, Irma...}. This fact forces them to appear in the scope of some particle that is sensitive to focus alternatives to be properly interpreted. According to Cable’s account, Q is a particle that is sensitive to the focus alternatives generated by wh-words. The semantic role of the Q particle is to convert the focus semantic value into an ordinary semantic value so that it can be interpreted by higher operators. In this sense, Q is sensitive to the focus value of its sister.

(53) Movement of Q Particle

While this analysis has empirical and conceptual advantages—particularly in the do-
main of pied-piping—at first blush it appears to move wh-movement and focus movement farther apart as syntactic phenomena. Under this analysis, wh-movement is necessarily indirect, due to a particular semantic property of wh-words which is not shared by foci. However, though foci do not share this semantic deficiency, like wh-words they generate focus alternatives. As I will argue in the following subsection, Q is not the only focus sensitive particle that can be targeted for attraction, in fact, Q particles are members of a larger class of focus sensitive particles that undergo Ā-movement (Cable, 2010; Branan and Erlewine, 2020).

4.2 Extension to Focus Movement

There is no consensus about how foci are displaced. Some previous work, noting the similarity of focus fronting to other Ā-movements, argues that foci bear a formal syntactic feature which can be targeted for attraction (Horvath, 1995; Frascarelli, 2000; Frascarelli and Puglielli, 2007a; Aboh, 2016). Often, this work assumes that focus fronting is triggered by a designated head in the left periphery (Bródy, 1990; Rizzi, 1997). This view has often been challenged, however, in the literature on displacement of foci. Specifically, much previous work has argued that a formal feature marking foci is empirically or conceptually undesirable, given that the focus of a sentence is determined with respect to the discourse context (Szendrői, 2001; Horvath, 2007; Fanselow, 2006, 2008; Reinhart, 2006; Chomsky et al., 2019). Instead, these authors primarily point to the unique prosodic character of foci or some aspect of their semantic interpretation to explain why they appear in non-canonical positions.

One important attempt to resolve these theoretical issues comes from Horvath (2007). On the basis of Hungarian, Horvath argues against the existence of a formal feature marking foci in the language. Instead, she argues that movement targets an operator which
c-commands some foci and which contributes an exhaustive interpretation. When this operator is targeted for movement, it moves the foci that it c-commands along with it.

(54)  Movement of Exhaustivity Operator

This analysis is remarkable similar in spirit to Cable’s analysis of wh-movement (a similarity that Cable also notes); both propose that the apparent movement of wh-words and foci actually takes place indirectly: wh-words and foci move only due to the fact that they stand in some structural relationship to a moved element. However, there is an important difference. Horvath argues explicitly, on the basis of Hungarian, that this movement is only sensitive to exhaustively interpreted elements, not foci per se. Recall from section 2.4 that only exhaustively interpreted foci undergo syntactic movement in Hungarian. However, given that wh-movement does not trigger an exhaustive interpretation in Hungarian (see discussion in section 2.4), it would be reasonable to conclude, on the basis of Hungarian alone, that wh-movement and “exhaustivity-movement” are two fundamentally distinct phenomena.

Horvath’s analysis of focus movement in Hungarian cannot be straightforwardly applied to SMPM. As shown in section 2.4, SMPM does not distinguish between exhaustive and non-exhaustive foci: both must be displaced to a preverbal position. Thus, the facts of SMPM allow us to view the puzzle from a slightly different angle, which invites us to
reconsider the possibility that wh-movement and focus movement are in fact connected.

I propose that all movement of wh-words and foci in SMPM results from attraction of a lexical class of particles that are sensitive to the focus semantic value of their c-command domain. In SMPM, this class includes a null Q particle that c-commands wh-words, overt focus particles such as *inta* (only) and *ntsya* (even), and a segmentally null particle that c-commands wh-congruence foci and contrastive foci.6 These particles share several important properties. First, they are all sensitive to the focus semantics of their sister. Thus, they all must c-command some element that introduces focus alternatives in order to be interpreted, similar to focus sensitive particles like English *only* (Jackendoff, 1972). Second, I propose that the particles in this class bear a formal syntactic feature which can trigger their attraction. For concreteness, I will refer to this feature as [*alt*], signifying the fact that all of these particles are sensitive to focus alternatives.7 In SMPM, these particles enter into an Agreement relationship with a head in the C-domain, triggering their attraction. A similar class of particles is independently proposed by Branan and Erlewine (2020) to account for patterns of pied-piping and “anti-pied-piping” cross-linguistically.

---

6While it doesn’t have any segmental realization, preliminary evidence suggests that this last particle bears a tonal specification which can trigger tonal effects at the right edge of fronted foci. For reasons of space, I do not include this evidence here, and assume for the purposes of this paper that this focus particle is covert.

7I refer to this feature as [*alt*] instead of the more common [*FOC*] because the two are conceptually different. Most often, the formal feature [*FOC*] is argued to be present on foci themselves, allowing them to be attracted directly. I use [*alt*] to signal the fact that it is the particle sensitive to alternatives that bears a feature and is attracted.
This extends Cable’s theory of wh-movement to another type of Ā-movements, an idea suggested, but not explored in depth, by Cable. Additionally, it generalizes Horvath’s exhaustivity operator proposal for Hungarian to account for a wider-range of focus-phenomena. At its core, this proposal claims that wh-movement and focus movement are two phenomena that form a natural class. I take this to be a welcome result, given the long literature establishing the syntactic and semantic similarities between them (Chomsky, 1977; Horvath, 1986; Croft, 1990; Rizzi, 1997; É. Kiss, 1998a; Jayaseelan, 2001; Aboh, 2007; Erlewine, 2018, a.o.). Wh-movement and focus movement are not identical phenomena, however. As I will discuss in section 4.4, the Wh-over-Focus Generalization forces us to conclude that there is something that makes Q particles unique among the class of focus sensitive particles. Before addressing this issue in more detail, I test three predictions of this account using overt focus sensitive particles in SMPM.

4.3 Three Predictions

The above account makes three predictions about movement of foci in SMPM. For the sake of clarity, I test these predictions with overt focus particles in the language. First, if focus sensitive particles are attracted, not foci themselves, then a focus should not be able to be attracted and strand a focus particle in situ. This is indeed the case—in SMPM
a focus sensitive particle cannot be stranded in situ.

\[(56)\]

\[\text{a. } *\text{U’un yibá yá shishi Pedro inta }\_
\text{no vegetables it,NEUT eats P. only}
\text{Intended: No, Pedro only eats vegetables.}\]

\[\text{b. } *\text{U’un źá maéstra źá ka’an inta }\_
\text{tu’un sá’á}
\text{no CL teacher she speaks only language Spanish}
\text{Intended: No, only the teacher speaks Spanish.}\]

Second, this account predicts that the moved constituent and the element semantically focused can be non-isomorphic. This is a welcome prediction, as there is a long literature establishing the fact that, cross-linguistically, there can be mismatches in the size of the element that undergoes focus movement and the semantic focus (Jackendoff, 1972; Horvath, 2007; Branan and Erlewine, 2020). Indeed, this was one of the principle empirical motivations for Horvath (2007) to propose that focus movement is indirect in Hungarian. Because the formal feature that enters into an agreement relationship is on the focus sensitive particle, not the semantically focused element, non-focused words should also be able to front along with the particle.

One way to test this is by looking at possessive DPs in the language. In SMPM, one strategy for possession is for the possessum to take a PP complement, the head of which selects the possessor.
Because *inta* can take this entire DP as its sister, movement of a focus sensitive particle should be able to front an entire possessive DP, even if only the possessor is semantically focused. This prediction is also confirmed, as shown in (58).

(58)  
\[\text{a. [Inta kárrò ñà’á } \textbf{Natalia} \text{ tun]} \text{ nitsivi } \quad \text{[only car } \text{poss } N. \text{ it.wood broke.down]}
\]

‘Only Natalia’s car broke down.’

\[\text{b. [Inta sè’e } \textbf{Maria} \text{ yá]} \text{ kú’u } \quad \text{[only child M. } \text{neut } \text{sick]}
\]

‘Only Maria’s child is sick.’

Third, any head that is not c-commanded by the focus sensitive particle should not front when the particle is attracted. Once again, we can test this prediction with possessive DPs. If the focus sensitive particle is instead merged as a complement of the possessive preposition and takes only the possessor as its sister, then we predict that fronting the entire possessive DP will be ungrammatical, as the possessum and preposition are no longer c-commanded by the focus-sensitive particle. Once again, this prediction is confirmed (59).

(59)  
\[\text{a. *Karro ñà’á } \text{inta } \textbf{Natalia} \text{ nitsivi}
\]

‘Car poss only N. broke.down’
Intended: Only Natalia’s car broke down.

It is important to note that (59) is not ungrammatical due to a restriction on the position of the focus sensitive particle. inta can intervene freely between a possessor and its possessum, as shown in (60). When it does, only the possessor will front, leaving the possessum in situ.

(60)   [Inta Natalia] nitsivi kárro ñà’ā __
      only N. broke.down car  POSS
‘Only Natalia’s car broke down.’

4.4 Syntactic Properties of Focus Sensitive Particles

Much work on the interaction between wh-words and foci has often assumed that wh-words are a specific type of focus. That is, wh-words, like foci, introduce focus alternatives, but they also have some property which distinguishes them from other foci. Several papers make the explicit claim that wh-words bear two formal features—a [FOC] feature and a [WH] feature—to formalize this intuition (e.g. Lee, 1999; Bošković, 2002; Sabel, 2000; Kim, 2006; Bocci et al., 2020).

However, as correctly pointed out by an anonymous reviewer, it is important to distinguish here between two interrelated notions: a semantic property of introducing focus alternatives and a prosodic property of non-canonical pronunciation. While wh-words—like foci—introduce alternatives, they are not always prosodically focused with a pitch accent. This is the case, for instance, in English, where wh-words do not receive primary sentential stress in wh-questions, in contrast to foci (Gunter, 1966; Culicover and Rochemont, 1983). Wh-words are prosodically focused, however, when they remain in-situ and are interpreted as an echo question (Erteschik-Shir, 1986; Beck and Reis, 2018). For this reason, it is important to keep separate the notion of introducing focus alterna-
tives and being prosodically focused. These notions are especially important to separate when we consider languages, like SMPM, which mark foci primarily syntactically. In this paper, I will focus on the shared semantic property of foci and wh-words and its syntactic consequences. In the terminology of (Erlewine, 2018), they are both “formally focused.” I will set aside the important question of how this property correlates with non-canonical prosody.

Here, I adopt basic intuition that wh-words are a subclass of foci, but argue that it applies at the level of focus sensitive particles. That is, there is a class of particles that are sensitive to focus alternatives and bear the formal feature [ALT]. Q particles, which bear the feature [ALT] by virtue of the fact that they are sensitive to focus alternatives, also bear a separate feature [Q]. Furthermore, I assume that these features are arranged in a feature geometry (cf. Harley and Ritter, 2002). That is, they are not independent of one another, but instead the feature [Q] entails the feature [ALT].

(61)  

<table>
<thead>
<tr>
<th>Focus Sensitive Particles</th>
<th>Q Particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ALT]</td>
<td>[ALT]</td>
</tr>
<tr>
<td></td>
<td>[Q]</td>
</tr>
</tbody>
</table>

Feature geometries have often been used in syntax and morphology to capture the intuition that syntactic features are not unstructured, but in fact can be related to one another in principled ways (see e.g., Harley and Ritter, 2002; Béjar, 2003; Béjar and Rezac, 2009; Foley and Toosarvandani to appear; Coon and Keine, to appear, a.o.) For instance, if we accept that the ϕ-features [SPEAKER] and [PARTICIPANT] are part of the grammar of a particular language, then we want our theory of features to capture the fact that all speakers are necessarily participants (Harley and Ritter, 2002; Béjar, 2003, a.o.). In a similar way, if we want to capture the intuition that Q particles are members of a class of focus
particles, but subset of that class which has unique properties, one way to represent this relationship is by positing that Q necessarily bears a superset of the features present on other focus sensitive particles.

The idea that Ā-features can be arranged in entailment relationships is not new. Early work by Starke (2001) argued that Relativized Minimality can be sensitive to subclasses of features, accounting for the fact that specific wh-words can extract out of wh-islands. This fact led Starke to propose a representational constraint on wh-movement which allows subclasses to extract across superclasses.

\[(62)\]

\[\begin{array}{c}
\text{a. } \star \alpha \ldots \alpha \beta \ldots \alpha \\
\text{b. } \alpha \beta \ldots \alpha \ldots \alpha \beta \\
\end{array}\]

Starke (2001): 8

Building on these ideas, Abels (2012) and Aravind (2017) argue for more articulated Ā-feature geometries that include a common feature \([\text{OP}]\) for wh-words and foci. However, in the feature geometry proposed by Abels (2012) wh-words and foci are not featurally distinct. According to the geometry proposed in Aravind (2017), foci and wh-words bear \([\text{FOC}]\) and \([\text{WH}]\) features, respectively, in addition to each bearing an \([\text{OP}]\) feature. Thus, both argue that wh-words and foci have an overlapping featural representation, however, neither proposes a detailed description of how probing works in the Ā-domain. This will be the aim of the next section.

### 4.5 Articulated Probing in the Ā-Domain

I assume that movement is the product of two component parts: Agreement and Internal Merge (Chomsky, 2001; Starke, 2001; van Urk, 2015, a.o.). Furthermore, I assume that Agreement is also split into two distinct operations: i) a search operation \(\text{MATCH}\) (which finds features the probe is relativized to) and ii) a valuation operation \(\text{VALUE}\) (which transfers features to the probe) (Chomsky, 2000). I assume that the search operation is strictly
local, but that a probe need not necessarily be valued by the most local goal which it has matched with (Boeckx and Jeong, 2004). Thus, I assume that a probe relativized to some feature will match with the most local syntactic constituent in its domain that is specified for that feature (cf. Attract Closest and Relativized Minimality Chomsky, 2000; Rizzi, 1990). However, as I will argue below, the goal that ultimately values the probe need not be the first goal that it has matched with. Instead, the goal that can most completely satisfy the probe’s needs will value it. Consequently, under certain circumstances, a probe can be valued by, and attract, a non-local goal.

I assume that Internal Merge is not inherently constrained by Attract Closest, and is a “free” operation which simply combines two syntactic objects (Chomsky, 2001). Internal Merge is, however, constrained by economy, and only combines objects that have entered into an Agree relationship with one another (Chomsky, 1995). For our purposes, I assume that Internal Merge can only apply to goals which have valued the probe, simply matching with the probe is not sufficient for movement. Thus, phrasal movement occurs when a copy of a phrase that has valued the probe is Internally Merged as a specifier of the probing head.

I propose that SMPM has a probe on C that is relativized to the feature geometry [u\_ALT-u\_Q]. When this probe initiates a search, it will find goals within its domain that

(63) **Valuation is a prerequisite for Internal Merge**: A goal must value the probe in order to be attracted by that probe.

8 As noted by an anonymous reviewer, in other languages, more than just foci and wh-words are attracted to the same position. In Hungarian, for instance, the position occupied by wh-words and foci is also occupied by negative existential quantifiers and negative adverbs. This is not the case in SMPM: recall from section 2.2 that both foci and wh-words surface in a position above fronted negative indefinites and manner adverbs. This suggests that there is cross-linguistic variation related to what class of elements are in complementary distribution with one another. Furthermore, this may suggest that in some languages, multiple probes can be “bundled” onto a single head, leading to complementarity, whereas in other languages, the same probes can be hosted on separate heads (see e.g., Hsu, 2017, and discussion in §4.6). This account would predict that the ordering of these bundled probes should be fixed, such that it matches the structurally
have matching features. Following Béjar (2003), I assume that any goal that has a feature that entails the root of the probe can match with it. Thus, all types of focus sensitive particles can potentially match with the probe, as all bear the \([\text{ALT}]\) feature that is at the root of the probe.

Match is an operation that is constrained by locality. That is, if there are two goals that could potentially match with the probe, it will first match with the most local goal (as defined by asymmetric c-command). In the case under investigation, this means that when a focus particle in subject position c-commands a Q particle in object position, the probe on C will first attempt to match with the focus sensitive particle in subject position. For presentational purposes, I represent the \text{MATCH} operation with a dashed arrow in the following trees.

\begin{itemize}
  \item[64] Focus > WH
  \begin{itemize}
    \item CP
    \item \(c\)
    \item \([u_{\text{ALT}}]\) \(v\)
    \item \([u_{\text{Q}}]\) \(\text{PartP}\)
    \item \([\text{ALT}]\) \([\text{ALT}]\) \([\text{Q}]\)
  \end{itemize}
\end{itemize}

Though this goal matches with one of the features of the probe, it is not a complete match. The \([u_{\text{Q}}]\) feature of C has not yet found a match, and thus cannot receive a valuation. I assume, following Béjar and Rezac (2009), that this can trigger a second search position of the probes in languages where they are not bundled. Given that in SMPM, fronted foci and wh-words are structurally higher than manner adverbs and negative indefinites, this account predicts that the probe that is looking for wh-words and foci should initiate a search before the probe that searches for negative existentials and negative manner adverbs in languages where these probes are bundled onto a single head. I leave to future work the investigation of this prediction.
cycle to attempt to find a match for \([u_0]\). That is, if any features on the probe remain unmatched after the probe has interacted with the first goal, these features can probe again, possibly finding a different goal within the clause to match with. This means that a single probe can match its features with both the subject and the object, as long as the subject does not exhaustively match its features (cf. Feature Gluttony, Coon and Keine, to appear). Concretely, this means that if the probe on C does not find a Q particle in subject position, it will search again, potentially finding one in object position. If there is a Q particle in object position, then the \([u_0]\) feature on the probe will match with it.

\[(65)\]  Focus > WH

\[
\begin{array}{c}
\text{CP} \\
\text{C} \\
[\text{u}_{\text{ALT}}] \\
\text{v} \\
\text{PartP} \\
[\text{u}_0] \\
\text{QP} \\
[\text{u}_{\text{ALT}}] \\
\text{Q} \\
\end{array}
\]

Thus, after two search operations, the probe has matched with two goals:

\[(66)\]  Set of Matched Goals (S): \(\{\text{PartP}_{\text{ALT}}, \text{QP}_{\text{ALT-WH}}\}\)

\(^{9}\)A crucial component of the system proposed in Béjar and Rezac (2009) is the notion that any unvalued features on the probe are reprojected to a structurally higher position. This is important for their analysis because they consider cases where the probe is located on \(vP\), and can only find the internal argument on the first cycle of Agreement. In the cases they consider, the external argument is only within the search domain of the probe once it reprojects (termed Cyclic Expansion). While there is nothing inconsistent about the concept of reprojection with the facts of SMPM, the configuration that I am considering is different. Specifically, I am considering cases where the probe is located on C can find the subject on the first search cycle. Because the probe is located on C, reprojecting the probe will not allow access to any goals that it did not already c-command. Lacking any empirical motivation for this aspect of their analysis, for simplicity sake I do not adopt it here. However, I do not intend to make any claims about whether features can reproject in general.
Because articulated probes have more than one feature that they need to match and value, it is possible that some goals will only partially satisfy their needs, while other goals will completely satisfy their needs. If an articulated probe bears an EPP feature (that is, requires something to move into its specifier), by what mechanism does it decide which of the multiple goals it has matched with will value it and be attracted? I propose that when an articulated probe requires something to be merged into its specifier and it has matched with multiple goals, the goal that more completely satisfies the needs of the probe will value it and be Internally Merged. Formally, I propose the constraint \textsc{Economize Valuation}, defined in (67) (cf. Oxford, 2014, 2019; Coon and Bale, 2014; van Urk, 2015):

\begin{equation}
\text{(67) \ \textsc{Economize Valuation}: A probe } P \text{ that has matched with a set of goals } S \text{ is valued by goal } G \text{ in } S \text{ such that the number of features on } G \text{ relevant to } P \text{ is greater than the number of features for any other } G' \text{ in } S.\end{equation}

At the core of this constraint is the notion that movement is triggered, and only happens to the extent necessary to converge a derivation. Because valuation of a probe with an EPP feature is “costly” (i.e., it triggers an Internal Merge operation), this is an operation this is, in principle, subject to economy. A similar idea is presented in Pesetsky and Torrego (2001):

\begin{equation}
\text{(68) \ \textbf{Economy Condition: A head } H \text{ triggers the minimum number of operations necessary to satisfy the properties (including EPP) of its uninterpretable features.}}\end{equation}

Pesetsky and Torrego (2001): 359

(67) is also inspired by, though slightly different, than several recent proposals (called “Best Match” in those papers) which attempt to account for patterns where an agreement morpheme will track either the subject or the object, depending on their feature specifications. To account for this, Coon and Bale (2014) and van Urk (2015) argue that a probe
can skip a local goal when a non-local goal would better satisfy the needs of the probe. Similarly, Oxford (2014, 2019) shows on the basis of Agreement morphology in Algonquin that a probe on Infl will agree with either the subject, the object, or both, depending on which is a better match for the needs of the probe. That is, the probe, will track the subject or the object, depending on which shares more of its features.

Recall that in derivations with no wh-words, foci move obligatorily to the left periphery. This fact can be captured using the same assumptions about the Agree mechanism. First, the probe matches with the focus particle. As the [Q] feature remains unmatched after this first cycle, it probes again attempting to find another match. However, as there is no goal that bears the matching feature within the probe’s domain, it will not match with any goal. I assume, following Preminger (2014), that it is necessary for a probe to initiate a search, but that failure to Agree does not crash the derivation.

\[(69)\]

After matching all the features that it can, the probe on C has entered into a matching relationship with the following set of goals:

\[(70)\] Set of Matched Constituents (S): \{PartP_{ALT}\}

According to ECONOMIZE VALUATION, the probe will be valued by the focus sensitive particle
that it has matched with. Though it doesn’t completely satisfy the probes needs, there is no other goal that the probe has matched with that can satisfy its needs better.

Importantly, the fact that foci move obligatorily in the absence of a wh-word demonstrates that focus sensitive particles are not simply invisible to the probe on C—they can be attracted just in case there are no Q particles in the derivation. An integral part of this system then, is the notion the valuation of a probe is delayed until all possible match operations have been exhausted within its domain. If valuation were immediate after match, then we would expect focus sensitive particles to be attracted as soon as they are matched, even if a subsequent match operation would find and attract a Q particle. Assuming that valuing a movement-triggering probe is sufficient to be internally merged into its specifier, then if a focus sensitive particle in subject position could immediately value the \([\text{u} \text{ALT}]\) feature as soon as it enters into a match relationship, we would expect it to be able to move regardless of whether a subsequent Match operation finds a Q Particle. Moreover, if the \([\text{u}Q]\) feature is subsequently valued by a Q Particle, then we would expect it to also be attracted, contrary to fact. In other words, assuming that all valuation of a movement-triggering probe results in internal merge, only the Q particle is valuing the probe in Wh-over-Focus derivations.

This tendency to delay and economize valuation may not be a universal property of the Agreement mechanism. According to (Coon et al., to appear), in Mayan languages, probes can be immediately valued by arguments that they match with, leading in some cases to a derivation crash. Specifically, the authors account for a restriction on \(\text{Ā}-\)extraction of ergative arguments by positing an articulated probe searching for \([\text{u}\text{Ā}, \text{u}\text{D}]\) features. The authors propose that when this probe matches with absolutive argument (which they argue moves to a structurally higher position than the ergative argument), its \([\text{uD}]\) feature is immediately valued. If the ergative argument bears an \([\text{Ā}]\) feature, then the probe will match with it and also be valued by it. Because each valuation must trigger a movement
operation, these two “conflicting” valuations from two separate goals will create irreconcilable demands on the probe, leading to a crash of the derivation. Specifically, the issue relates to the timing of movement: both goals need to move, but they can’t move simultaneously or sequentially. Following (Coon and Keine, to appear), the authors refer to this pattern as “Feature Gluttony.”

(71)  “Feature Gluttony” in Ā-Probing

[Coon et al., to appear]: 18

Thus, if the account in Coon et al. (to appear) is correct, then some languages may not economize over the valuation operation in the same way that SMPM does. However, it is important to note that straightforwardly adopting the approach of Coon et al. (to appear) would incorrectly predict that a wh-word could not move across a more local focus in SMPM, given that Q particles bear a superset of the features of other focus sensitive particles. Thus, I propose that in SMPM this derivation crash can be avoided by delaying valuation until all match relationships have been established.

4.6 Considering an Alternative

In the previous subsection, I argued that there is a single probe on C in SMPM which searches for both [ALT] and [Q] features. In order to account for the Wh-over-Focus Generalization, I argued the the grammar economizes valuation, ensuring that a Q-particle will value the probe if it is able to. In this subsection, I briefly consider an alternative approach to bundled Ā-probing.10

Some previous work has argued that multiple Ā-features can be bundled on a single head. This approach, for instance, is taken by Hsu (2017) to account for languages

10I thank an anonymous reviewer for suggesting the possibility of this approach.
that have “relaxed” V2 requirements, and by Kotek (2014) to account for the interaction between superiority violating multiple wh-questions and focus intervention effects in Hebrew. In particular, Hsu (2017) adopts the “Feature-Scattering Hypothesis” of Giorgi and Pianesi (1997): in some languages, multiple Ā-features are bundled on a single head in the left periphery, while in other languages, these same features can be “scattered” across several distinct heads.

Kotek (2014) assumes a similar bundling of features on a single head to account for cases where superiority-violating multiple wh-questions in Hebrew do not trigger Focus Intervention Effects. Abstracting away from the details, she posits a head that bears both a [uWH] probe and a [uQ] probe. Furthermore, she argues that Agreement can be parasitic: when a probe establishes an Agreement relationship with some goal, all other probes on the same head can attempt to establish Agreement with that goal. Kotek (2014) assumes that the order of probing of the [uWH] feature and the [uQ] is not fixed. In principle, either can probe first, and in fact this fact allows her to account for the fact that only some superiority-violation questions in Hebrew display Focus Intervention Effects.

Thus, an alternative account for the Wh-over-Focus Generalization in Mixtec presents itself: [ALT] and [Q] are features bundled on the same probe. When [Q] probes first, it can find a Q particle in object position, ignoring any other structurally higher focus particles. Then, by Parasitic Agreement, the [ALT] probe can also enter into Agreement with the Q particle, bypassing the structurally higher focus particle completely.

In order to evaluate this proposal, it first necessary to establish a key assumption that we would be forced to make. To account for the pattern of Mixtec, the [Q] feature would have to probe before the [ALT] feature. If [ALT] is allowed to probe first, then we would predict a derivation where a focus particle in subject position enters into an Agreement with the [ALT] probe, then the [Q] probe searches for and finds a Q particle in object position. Given that focus sensitive particles move in the absence of a wh-word, it is safe
to assume that valuation of the $[u_{\text{ALT}}]$ feature alone can trigger movement to its specifier. Thus, allowing $[\text{ALT}]$ to probe first would make an incorrect prediction. If $[u_Q]$ has an independent EPP feature, then we predict that a focus over WH configuration will trigger movement of both. If, on the other hand, $[u_Q]$ does not have an independent EPP feature, then we predict that a focus sensitive particle in subject position will be attracted, leaving a Q particle in situ. Given that neither of these possibilities are possible in SMPM, this alternative analysis would require the stipulation that the $[Q]$ feature always must probe first. Furthermore, a fixed order of probing is required to capture the cross-linguistic Wh-over-Focus Generalization. Supposing that other languages (besides SMPM) allowed for $[u_{\text{ALT}}]$ to probe first, then we would expect to find languages that move either a focus or a wh-word, whichever is structurally higher. To my knowledge, this pattern of movement has not been attested.

Thus, in contrast with the assumption that Kotek (2014) makes about articulated probing in Hebrew, this alternative analysis for the Wh-over-Focus Generalization would be forced to state that $[Q]$ probes before $[\text{ALT}]$ cross-linguistically. Consequently, when evaluating this alternative hypothesis, we should ask: is there an empirical or theoretical motivation to stipulate that $[u_Q]$ must always probe before $[u_{\text{ALT}}]$?

In his theory of bundled probes, Hsu (2017) proposes that the ordering of probing is not variable, but fixed according to a universal ordering constraint (Giorgi and Pianesi, 1997). This constraint ensures that there will be a correlation between languages that realize several Ā-features on a single head, and those that realize the features on distinct heads within the C-domain.

(72) Universal Ordering Constraint:

Features are ordered so that given $F_1 > F_2$, the checking of $F_1$ precedes the checking of $F_2$. 

54
In short, this constraint states that if two features are realized on two distinct heads in one language (with $F_1$ always structurally higher than $F_2$), then when those two features are bundled together on the same head in another language, $F_1$ will probe before $F_2$.

Thus, following the empirical motivation for fixed probing leads us to the following prediction: if [uo] is universally ordered to precede [ALT], then we expect that fronted wh-words to always precede fronted foci in languages where both can move. In other words, this fixed order of probing should correlate with fixed structurally positions within the left periphery.

Indeed, in some languages, wh-phrases must precede focused elements when they both move to the left-periphery. In Toba Batak (Austronesian), for instance, when both a wh-phrase and a focus are fronted, the wh-word will precede the focus.

(73) Ise holan indahan di-allang ___?
who only rice PASS-eat
‘Who ate only rice?’

To be empirically consistent with the proposal in Universal Ordering Condition proposed in Giorgi and Pianesi (1997), we would expect that this ordering should be universal. However, in other languages, such as Babine-Witsuwit’en (Athabaskan), fronted foci must precede fronted wh-words.

(74) Hoo’, lhēs ʼiy nts’ē Lillian ___ yunkēt
no bread FOC where L. 3sg.bought.3sg
‘No, where did Lillian buy the bread?’ (not the fish)

Furthermore, according to Erlewine (2018), the fixed order of Wh over focus in Toba Batak is driven by semantic, rather than syntactic factors. He argues (fn. 11, pg. 668) that the opposite order (Focus > WH) is ruled out by a semantic Focus Intervention Effect.
Thus, a fixed order of Wh > Focus in languages where they both move does not seem to be empirically or typologically motivated. Consequently, there is no clear motivation to posit a universally fixed order of probing of [uQ] > [uALT].

Finally, it is worth considering the conceptual differences between this approach and the analysis proposed in the previous subsection. If the order of probing is fixed, then there must be some independent principle that restricts lexical entries cross-linguistically, such that the head in (75a) is allowed, but heads of the form (75b) are disallowed (where the order of heads reflects the order of probing, from top to bottom).

\[
\begin{align*}
(75) \quad & \text{a.} \\
& \begin{array}{c}
H^\circ \\
\phantom{uQ} \\
\phantom{uALT}
\end{array} \\
& \begin{array}{c}
\phantom{H^\circ} \\
\phantom{uQ} \\
\phantom{uALT}
\end{array}
\end{align*}
\]

\[
\begin{align*}
(75) \quad & \text{b.} \\
& \begin{array}{c}
\phantom{H^\circ} \\
\phantom{uQ} \\
\phantom{uALT}
\end{array} \\
& \begin{array}{c}
H^\circ \\
\phantom{uALT} \\
\phantom{uQ}
\end{array}
\end{align*}
\]

It does not seem clear to me why such a restriction on lexical items would be warranted. It is not so simple as to say that the head in (75b) is unlearnable. In fact, a hypothetical language with that lexical item would display a unique pattern of fronting which would presumably give a child enough information to posit it. Specifically, the head in (75b) would move the structurally highest focus if there are two in the derivation. If a wh-word is structurally higher than a focus, then it would move alone. If, however, a focus was structurally higher than a wh-word, then they would both move. Thus, a language with such a probe would display a complex pattern of fronting related to the structural positions of wh-words and foci. Given that children are often forced to make inferences about lexical items on the basis of complex patterns, I see no reason why this pattern would be particularly unlearnable.

In contrast, my proposal places this cross-linguistic restriction not in the lexicon, but in the syntax proper. It is optimization in probing and valuation that triggers the WH-over-Focus Generalization, not optimization of lexical items across languages.
5 Focus Intervention and the Identity Hypothesis

Recall that in §2, I considered and rejected the Identity Hypothesis, which proposes that foci and wh-words are formally identical in the syntax. In part, I rejected that hypothesis because, assuming standard syntactic locality, it predicts that in languages where foci and wh-words are in competition for movement, the structurally highest one will move. In fact, evidence from several languages, including SMPM, suggests that in clauses where wh-words and foci co-occur, a wh-word will front across a focus (76a). Fronting a structurally higher focus, predicted to be possible by the Identity Hypothesis, is impossible (76b).

(76)

a. Object Wh-Word Moves, Subject Focus Remains in situ

\[ \text{WH ... FOC} \]

b. Subject Focus Moves, Object Wh-Word remains in situ

\[ \text{*FOC ... WH} \]

In this section, I consider whether it is possible to maintain the Identity Hypothesis by identifying an independent constraint that would rule out the configuration in (76b). One reasonable hypothesis is that (76b) is impossible because it triggers a Focus Intervention Effect, sometimes called a “Beck Effect.” (Beck, 1996, 2006; Hagstrom, 1998; Pesetsky, 2000; Kim, 2002; Li and Law, 2016; Branan, 2018; Kotek, 2019, a.o.). Broadly speaking, in many languages, there is a surface representational constraint against a focus or quantificational element intervening between a wh-word in situ and its licensing complementizer. For example, in Korean, wh-words are normally left in situ (77a), however, when a focus sensitive operator like only c-commands the wh-word, the sentence becomes ungrammatical (77b). In order to express the meaning intended in (77b), the wh-word must scramble to a position where it is no longer c-commanded by the focus sensitive operator.
This effect has been demonstrated for a wide range of wh-in situ languages, and has also been claimed to hold in some circumstances when wh-words are left in situ in languages that normally move them (Beck, 2006; Pesetsky, 2000). In particular, Pesetsky (2000) demonstrates that intervention effects arise in English in superiority-violating multiple wh-questions, but not in superiority-obeying wh-questions (see also, Kotek, 2019).

It is worth considering, then, whether (78) is ungrammatical in SMPM because a focus intervenes between the wh-word in situ and its licensing complementizer.

(78)  
*Pedro rà kishashi __ nà  
P. he brought what  
Intended: What did Pedro bring?

Under this analysis, movement of a focus instead of a wh-word would be impossible because the resulting derivation would produce a marked structure: namely, a wh-word in situ that is preceded by a focus. Consequently, an advocate of this analysis might argue that the Identity Hypothesis can be maintained, as the difference between foci and wh-words is their semantic properties, not their syntactic features; the constraint against syntactic movement of local foci is explainable by other means.
5.1 Background on Focus Intervention

While there have been many attempts to explain focus intervention, much recent work has adopted the basic proposal of Beck (2006): Focus Intervention arises when a wh-word is in the scope of a focus sensitive operator (see also, Cable, 2010; Kotek, 2019).

(79) **Focus Intervention:**  
\[ Q \ldots [O \phi \ldots X P_F \ldots wh \ldots]] \]

Beck’s analysis relies on the notions of ordinary semantic value and focus semantic value. The ordinary semantic value of a phrase (notated \([\alpha]^o\)) is its denotation, and its focus semantic value (notated \([\alpha]^f\)) is a set of alternatives that is generated by replacing any focus generating element within that phrase with anything of the same semantic type (Rooth, 1992). Thus, the ordinary semantic value of a phrase like \([American]_F\text{ farmer}\) is a function which maps some individual \(x\) to the proposition that \(x\) is both American and a farmer (80a). The focus semantic value of the same phrase is a set of propositions, stating that \(x\) is a farmer, and that is has some other property \(P\).

(80) a. **Ordinary Semantic value of \([American]_F\text{ farmer}\):**  
\[ \lambda x [\text{American}(x) \land \text{farmer}(x)] \]
b. **Focus Semantic Value of \([American]_F\text{ farmer}\):**  
\[ \{\lambda x [P(x) \land \text{farmer}(x)]\} | P : E \rightarrow propositions \]

Beck (2006) proposes that wh-words have a focus semantic value, but do not have an ordinary semantic value. That is, they introduce alternatives, but do not have a denotation. In normal circumstances, they can be interpreted when they appear in the scope of a question operator, which converts the focus semantic value of its scope directly to an ordinary semantic value. Thus, a question like (81a) has the focus semantic value in (81b), which is directly converted by the question operator into an ordinary semantic value. This
is consistent with the hypothesis that the meaning of a question is the set of propositions which could answer that question (Hamblin, 1973; Karttunen, 1977).

(81)  a. Who makes the best mole in Oaxaca?
     b. {Teresa makes the best mole in Oaxaca, Vitorino makes the best mole in Oaxaca, Gloria makes the best mole in Oaxaca...}

Thus, the question operator semantically composes with the focus semantic value of its sister, but does not make any reference to its ordinary semantic value. Consequently, a wh-word, which has no ordinary semantic value, can freely occur within the scope of a question operator.

Other focus operators (besides the question operator), use both the focus semantic value and the ordinary semantic value of their scope as part of their semantic composition (Beck, 2006; Cable, 2010). In the process of normal focus interpretation, the focus operator “uses up” the focus alternatives of its scope and returns the ordinary semantic value of its scope for use by higher operators. In other words, focus operators require that their scope have a well-defined ordinary semantic value that can be returned. Because wh-words do not have an ordinary semantic value, when they appear in the scope of a focus operator that scope will have an undefined ordinary semantic value, creating an uninterpretable structure. Assuming that every semantic derivation must have a defined ordinary semantic value (Beck’s Principle of Interpretability, 16), then once a focus operator has acted upon the wh-word, there is no way to salvage the derivation.

Furthermore, even if the focus operator were somehow able to interpret a wh-word in its scope, it would still disrupt the projection of focus alternatives introduced by the wh-word to the Q operator (Kotek, 2019). Because focus operators “reset” the focus semantic value of their scope to an ordinary semantic value, this means that there would then be no focus semantic value for the Q operator to interpret. Without a focus semantic value, the
Q operator is not able to generate the right interpretation of a question as a set of propositions.

Some previous work has tried to account for Focus Intervention Effects syntactically by arguing that certain intervenors block covert movement of in-situ wh-phrases (Beck, 1996; Pesetsky, 2000). However, Kotek and Erlewine (2016) and Kotek (2019) argue convincingly that covert movement across an intervener does not trigger a FIE. In fact, covert movement across an intervener is used as a strategy to avoid a FIE, as covert movement allows wh-words to scope outside of focus operators. The crucial evidence for this comes from a series of observations about the relationship between the possibility for covert movement and the presence of intervention effects. Kotek (2019) shows that, in English, the possibility for covert movement correlates with a lack of intervention. However, when covert movement is blocked (for instance, when a wh-phrase contains an NPI that must be licensed, a focus that must be associated with, or an anaphor that must be bound) intervention effects emerge. When there is no restriction on covert movement, there are no intervention effects. Consequently, Kotek argues that intervenors do not block covert movement, but rather, covert movement is a strategy to avoid a semantic intervention effect. Thus, she argues that FIE must be an LF phenomenon, rather than a syntactic or prosodic one.

5.2 Determining the Scope of the Focus Operator in SMPM

According to the semantic account of FIE introduced in the previous subsection, ungrammaticality should only arise when the wh-word is in the scope of the focus operator. If the focus takes narrow scope (not including the wh-word), then it can return an ordinary semantic value for its scope without problem. Furthermore, if the focus operator doesn’t c-command the wh-word, then it will not block the alternatives introduced by the wh-
word from being interpreted by the Q operator—the alternatives of the wh-word will not be used up by the focus operator and therefore will remain accessible to the question operator.

(82) **No Focus Intervention:**

\[ [Q\ldots[\text{OP }\phi\ldots X P_F]\ldots w h\ldots] \]  

(cf. 79)

So, in order to determine if the Wh-over-Focus Preference is reducible to a FIE, we first need to establish the scope of the focus operator in SMPM. If the wh-phrase must be within the scope of a focus operator, then the ungrammaticality of this configuration could plausibly be attributed to an FIE. However, if we can demonstrate that moving a focus instead of a wh-word is ungrammatical *even in the case* that the wh-phrase is outside the scope of the focus operator, then we will have good evidence to suggest that the Wh-over-Focus Preference is not reducible to a FIE.

In SMPM, contrastive foci can take narrow, DP-level scope, as they do in English. However, even in cases when a subject in focus is interpreted contrastively, it still cannot move instead of a wh-word. I assume, following Rooth (1992), that contrastive foci are interpreted using the ~ operator. This operator introduces a variable, which can be coindexed with other semantic objects. However, only semantic objects whose ordinary semantic value is part of the set of alternatives of the phrase within the scope of the focus operator can be coindexed with the variable introduced by the focus operator. Put differently, this restriction forces the variable to be coindexed with a phrase that contrasts with the phrase in the scope of the focus operator.

(83) **Contrasting Phrases:** Construe a phrase $\alpha$ as contrasting with a phrase $\beta$, if $[\beta]^o \in [\alpha]^f$.  

Rooth (1992): 81

In a famous example, Rooth shows how *American farmer* and *Canadian farmer* are
contrasted with one another in ((84)).

(84) An ~[**American** farmer] met a ~[**Canadian** farmer].

Rooth (1992): 86

A ~ operator takes *American farmer* in its scope, and introduces a variable that could be contrasted with anything that is part of its alternative set, e.g. Canadian farmer, Mexican farmer, French farmer, etc. A second ~ operator introduces another variable that can be coindexed with anything that is part of the alternative set of *its* scope, e.g. American farmer, Mexican farmer, French farmer, etc. Because each type of farmer is part of the alternative set of the other, they can be interpreted contrastively.

Because of the way that alternative sets are generated, altering the scope of the focus operator will change the set of phrases that can be construed contrastively with the focus. (84) shows that in English, contrastive foci can take narrow scope. If the operator took sentential scope, then the alternatives generated would be of the shape *A x farmer met a y farmer*. In other words, we would expect (84) to contrast with another meeting event, such as between a Mexican farmer and a Guatemalan farmer. Intuitively, however, this is not the interpretation. Instead, it is meant to convey a contrast between the American and the Canadian. In short, if the focus operator takes DP level scope, then DPs can be contrasted with one another (as in 84). If, however, it takes sentential level scope, then the sentence will only contrast with another sentence.

In SMPM, fronted contrastive foci can be contrasted with other DPs, analogously to contrastive foci in English, as shown in (85).

(85) a. ~[**Tsina** tohǔ rí] ntsíku ~[**tsina** yā]
   dog black **AML** chases dog white
   ‘The *black* dog is chasing the *white* dog.’

   b. ~[**Kárro** ñà’ã **Pedro** tūn] kama chága nuhū ~[**kárro** ñà’ã **Juan**]
   car POSS P. it. **WOOD** fast more than car POSS J.
   ‘**Pedro**’s car is faster than **Juan**’s car.’
In each of these two examples, the natural interpretation is one of contrast between two DPs. In (85a), two different dogs are contrasted with one another, and in (85b) two different cars are contrasted with one another. If the fronted focus took sentential scope, then instead of this interpretation, we would expect (85a) to contrast with other chasing events between two dogs, and (85b) to contrast with other speed comparisons between two cars. Given that this is not the interpretation of this sentences, this suggests that the ~ operator can take DP-level scope in the language.

With this fact in mind, consider again the example in (86). Given the context, it is clear that the fronted focus Pedro is interpreted contrastively with respect to the salient alternative Pablo. Consequently, the wh-word is not in the scope of the focus operator, which is taking DP-level scope. However, as we have seen, this configuration is still ungrammatical.

(86) Context: My friend Benjamín and I went a party where everyone brought some food or drink to share. We know what Pablo brought, but we didn’t see what Pedro brought. When we go to ask the host, Benjamín mistakenly asks him, “What did Pablo bring?” I turn to Benjamín and say, No:

   a. *~[Pedro rã] kishashi __ nã
      P. he brought what
      Intended: What did Pedro bring?

Thus, straightforwardly adopting the analysis proposed in Beck cannot fully explain the Wh-over-Focus Preference. According to Beck’s analysis, (86) should not trigger a Focus Intervention effect, and thus the fact that it is ungrammatical must be explained by other means. In other words, the Wh-over-Focus Preference is not reducible to a semantic constraint of this type.

Of course, this doesn’t preclude the possibility that some instances of moving a focus instead of a wh-word could trigger a FIE. For instance, recall that focus sensitive opera-
tors like the equivalent of *only* also move to a preverbal position in SMPM.

(87) \[\text{Inta } \text{Maria }\ddot{\text{ñ}}\ddot{\text{a}} \text{ và’a kása’a rá} \]
only Maria she well makes it.LIQ
\[\text{‘Only Maria can make it (mole).’}\]

If *inta* has similar semantics to English *only*, then the alternatives it operates over are the focus semantic value of the entire sentence (Kotek, 2019).

(88) \{\text{Maria can make it, Gloria can make it, Juan can make it}\}

If this is correct, then in (89), the wh-word is in the scope of the focus operator. In other words, (89) is ungrammatical for multiple reasons: (1) it is ungrammatical because a focus has moved instead of a wh-word (in violation of ECONOMIZE VALUATION); (2) a wh-word is inside the scope of a focus operator, triggering a focus intervention effect.\textsuperscript{11}

(89) \[\text{Context: There is some mushroom in the forest that nobody likes the taste of except Maria.}\]

a. \[\text{*Inta } \text{Maria }\ddot{\text{sháshi}} \_ \text{ntsyâ }\text{shí’ï?} \]
only M. eats which mushroom
\[\text{Intended: Which mushroom does only Maria eat?}\]

To summarize: there are instances where a focus moving instead of a wh-word will create a structure that is semantically uninterpretable, commonly called a Focus Intervention Effect. However, this marked structure only arises when a wh-word is in the scope of a focus operator. As I have demonstrated in this section, at least some focus operators can take narrow DP-level scope in SMPM. Consequently, the Wh-over-Focus Preference cannot be reduced to focus intervention alone and the challenge to the Identity Hypothesis remains.

\textsuperscript{11}I thank an anonymous reviewer for raising this point.
6 Conclusion

In this paper, I have considered several possible formal representations of wh-words and foci. First, I considered the possibility that they might be formally identical. The predictions of this hypothesis, however, are not borne out. In several languages, including San Martín Peras Mixtec, wh-words will always move instead of a focus, even if they are non-local. Second, I considered the possibility that they might have disjoint features, each being attracted by a separate head. This analysis faces empirical problems within SMPM, and also provides no explanation for why wh-words and foci should be attracted to the same apparent position in so many languages.

Instead, following some previous work, I proposed that a class of focus sensitive particles are attracted in SMPM: Q particles, which bear $\{\text{ALT}\}$ and $\{\text{Q}\}$ features, and other focus sensitive particles which only bear $\{\text{ALT}\}$. Additionally, I proposed a feature geometry which allows us to maintain the generalization that Q particles and other focus sensitive particles are attracted by the same head, while giving us a means of explaining the preference for wh-movement over focus movement. By adopting and expanding on two proposals from the A-domain—multiple searches and ECONOMIZE VALUATION—I proposed that an articulated probe could match with both the subject and the object under certain configurations. Furthermore, I claimed that the attracting head would always prefer to move a Q particle rather than another focus sensitive particle, due to the fact that it has a superset of the features. In this way, I showed that a feature geometric analysis can account for the apparent non-local movements of wh-words.

If this analysis is on the right track, then we may find analogies of hierarchy-based phenomena in the Ā-domain. In a sense, ECONOMIZE VALUATION captures the intuition that a head may prefer to attract a constituent that is higher on some hierarchy, just as some agreement slots preferentially agree with constituents that are higher on some hierar-
chy. We might, for example, expect to find languages that ban certain configurations of wh-words and focus, analogous to a PCC effect (e.g. Perlmutter, 1971; Bonet, 1991), or languages which use special morphology to mark the structural relationship between wh-words and focus, analogously to inverse morphology in Algonquian (e.g. Oxford, 2019; Aissen, 1997). Going forward, more needs to be done to investigate hierarchy effects in the Ā-domain. The preference for wh-movement over focus movement may be just one of many ways in which feature geometries in the Ā-domain manifest themselves.

References


de Hollenbach, E. E. (2013). *Gramática del mixteco de Magdalena Peñasco* (sa’an ſuu ſuu...


