New Information and the Grammar of Focus: 
Evidence from San Martín Peras Mixtec

Andrew Hedding

3/9/19

1 Introduction

A well-known observation: answers to wh-questions are prosodically prominent in English (Selkirk 1995, Jackendoff 1972, Rochemont 1986, a.m.o.)

(1) 
Q: Who bought a book about bats?  
A: Mary bought a book about bats.

Selkirk (1995):554

In this talk, I will explore similar prominence in San Martín Peras Mixtec (SMPM), a tonal language.

(2)  
a. Nā tāshĩ Mariana ntà’â Bernardo
what give.comp M. hand B.
‘What did Mariana give Bernardo?’

b. Chíchí[tāshĩ=ñá] ntà’â=rà
avocado give.comp=3sg.f hand=3sg.m
‘She gave him an AVOCADO.’

I will argue that answers to wh-questions (2b) display asymmetric prominence compared to answers to other types of questions.

– Word final high tones are raised in pitch but other tones are unaffected.¹

– Additionally, they undergo syntactic movement to a preverbal position which can be distinguished from other movement operations in the language.

I’ll argue that SMPM’s properties speak to an ongoing debate about the grammatical representation of answers to wh-questions.

¹Thanks first and foremost to Natalia Gracida Cruz for her patience and insights during the elicitation process, as well as for sharing her intuitions about her language with me. Many thanks are also due to Ryan Bennett, Pranav Anand and Gorka Elordieta, as well as Andrew Angeles, Jérémy Beauchamp, Richard Bibbs, Ben Eischens, René Gutiérrez, Lisa Hofmann, Ur Shlonsky, Maziar Toosarvandani, Marta Wierzba, an audience at the AIS Group at UCLA, and the members of the 290 Research seminar for very helpful comments and other help with the development of this project.

¹Here and throughout, a grave accent (̃) marks low tones, an acute accent (˘) marks high tones, a caron (˚) marks rising tone, and mid tones are unmarked. Pitch raising is represented with an arrow (↑). The following abbreviations are used in the glosses: comp=Completive, cop=Copula, f=Feminine, m=Masculine, sg=Singular.
Feature hypothesis: The answer is a type of focus, which is marked with a morphosyntactic feature (e.g. [foc]). This feature may:

- Directly trigger phonological prominence at PF (Selkirk 1995, Rooth 1992, Féry and Samek-Lodovici 2006);
- Indirectly trigger prominence via alignment to a prosodic boundary (Büring 2013, Féry 2013, a.o.).

(3) Q: Who bought a book about bats?
A: [mary]_{foc} bought a book about bats.

Givenness hypothesis: The answer is prominent because it is not given.

- Given (i.e. entailed by the context) constituents resist prominence (Schwarzchild 1999, a.m.o.)
- The non-given constituent receives default phrasal prominence (Kratzer and Selkirk 2018, Umbach 2009, Féry 2007).
- Wh-question answers are not represented formally in the grammar. (Kratzer and Selkirk 2018, Schwarzchild 1999)

(4) Q: Who bought a book about bats?
A: mary [bought a book about bats.]_{given}

In English, it can be hard to tell these two hypotheses apart, but today I will use the pattern in SMPM to argue in favor of the feature hypothesis.

The claim: Answers to wh-questions in SMPM are represented formally in the grammar.

The ingredients:

1. Answers to wh-questions are realized with prosodic prominence.
2. The effect is not caused by default phrasal prominence.
3. Answers to wh-questions are treated differently by the syntax than other constituents.

I will argue that these facts support the feature hypothesis—an account based solely on givenness struggles to account for the empirical generalizations.

Roadmap

- §2: I provide some relevant background on SMPM.
- §3: I detail the empirical generalization of wh-question answer prominence in the language.
- §4: I argue that this prominence is not default phrasal prominence.
- §5: I present two syntactic diagnostics that distinguish wh-question answers from other types of constituent fronting in the language.
- §6 concludes and presents future directions.
2 Background

San Martín Peras Mixtec is an Oto-Manguean language spoken in western Oaxaca, Mexico, near the coast of Guerrero.

Additionally, it is spoken by a sizable diaspora population throughout California.

— All data in this talk come from a single native speaker who lives and works in Watsonville, California.

SMPM is a VSO language, like other Mixtec languages (Macaulay 2005, Ostrove 2018). **Answers to wh-questions are fronted to a preverbal position.**

(5) a. \(Ná \ tā̄shĩ \ Ana \ ntā̄'ǎ \ Juan\)
   \(\text{what give.comp A. hand J.}\)
   ‘What did Ana give Juan?’

b. \(Shì'ǐ \ tā̄shĩ=ñā \ ntā̄'ǎ=rā\)
   \(\text{mushroom give.comp=3SG.F hand=3SG.M}\)
   ‘She gave him a mushroom.’

SMPM also allows fronting of constituents when responding to a broad focus (i.e. all new) triggering question.

(6) a. \(Ná \ kù \ bitsi\)
   \(\text{what cop.comp today}\)
   ‘What happened today?’

b. \(Shì'ǐ \ tāshĩ \ Ana \ ntā̄'ǎ \ Juan\)
   \(\text{mushroom give.comp A. hand J.}\)
   ‘Ana gave a mushroom to Juan.’

This allows for a **direct comparison of prosody**, without influence from declination (cf. DiCanio et al 2018).

SMPM is tonal, like all other Oto-Manguean languages (DiCanio and Bennett to appear).

— **It has 5 tones—low, mid, high, falling and rising** (Peters 2017).

  — Here I will not be discussing falling tones, which occur less frequently in the language.
  — All rising tones I consider are word final, which are mostly restricted to that position (Peters 2017).
Below, I give time-normalized plot of all target tones in initial position (Figure 1) and final position (Figure 2) across a variety of sentence contexts.

— Each vowel was isolated in Praat, and 10 pitch measurements were taken at equal points across the whole vowel in Equivalent Rectangular Bandwidth (ERB).\(^2\)

— Colored lines represent the time-normalized contour across the entire vowel (excluding the last measurement where there is a clear effect of tonal co-articulation).

— Gray bars represent a 95% confidence interval around the mean.

3 Answers to Wh-Questions are Prosodically Prominent

3.1 Methodology

— 36 target words were elicited 8 times as answers to wh-questions and as answers to a broad focus triggering question over the course of 10 elicitation sessions (n=576).

  - All words are bisyllabic roots.\(^3\)
  - All words the direct object in a ditransitive construction (as in 5 and 6).
  - Elicited using pictures and the questions in (7a) and (8a).

---

\(^2\)ERB is a psychoacoustic measurement scale that linearly plots the way pitch is perceived. That is, a difference of 1 ERB will be perceived as the same interval at any pitch. This is not the case with Hertz, where differences are perceived logarithmically (e.g. the difference between 220Hz and 440Hz on the one hand, and 440Hz and 880Hz on the other are heard as the same interval.)

\(^3\)Roots in Mixtec are minimally bisyllabic or bimoraic (Pike 1948, Macaulay and Salmons 1995, a.m.o.)
3.2 Results

The results of this systematic elicitation point to an asymmetry across tone types.

Word-final high tones are significantly raised when they are answers to wh-questions, when compared to broad focus contexts (Figure 5).\(^4\)

\[^4\]P-values are the results of T-tests where the alternative hypothesis is that the difference in means is greater than 0.02 ERB, the Just Noticeable Difference threshold. N-values represent the total number of syllables in each context. Differences are averages of pitch values across the syllable.
— This difference is **roughly half the distance between the onset of each of the level tones** (cf. Figures 1 and 2)

— This difference is **well above the Just Noticeable Difference (JND) threshold** of ≈0.02 ERB (Pulkki and Karjalainen 2015).

Unlike other tone languages which show pitch range expansion (e.g. Mandarin, Xu 1999), **low tones do not lower in pitch when they are part of an answer to a wh-question**.

Similarly, **mid tones are not affected** when they are an answer to a wh-question.
Rising tones also show no significant difference.

\[ \text{Ex: ya } \text{á} (\text{chile'}) \]

![Figure 10: Rising Tones](image)

\( n=191, \text{n.s.} \)

\[ \text{Difference}=0.05 \text{ ERB} \]

In addition to pitch, vowel duration is often a correlate of prosodic prominence (Ladd 2008, Gussenhoven 2004, Jun 2005).

In SMPM, **there is no significant difference in duration between the question contexts.**

![Figure 11: Syllable Duration across Tone Types](image)

These results suggest that **focus prominence is realized asymmetrically across tone type in SMPM.**

- Only high tones are realized at a higher pitch when they are answers to wh-questions.
- There is no duration difference across contexts.
Does Mixtec have Default Phrasal Prominence?

I argued in the previous section that SMPM realizes prominence on answers to wh-questions as raising of final high tones.

(9) **Wh-question**
   a. *Ná tâshi Mariana ntâ’â Bernardo*
      what give.comp M. hand B.
      ‘What did Mariana give Bernardo?’
   b. *Chichi tâshi=ñá ntâ’â=râ*
      avocado give.comp=3SG.F hand=3SG.M
      ‘She gave him a AVOCADO.’

(10) **Broad Focus Question**
   a. *Ná kù bitsi*
      what cop.comp today
      ‘What happened today?’
   b. *Chichi tâshi Mariana ntâ’â Bernardo*
      avocado give.comp M. hand B.
      ‘Mariana gave Bernardo an avocado.’

Recall from §1, there are two ways of thinking about this prominence:

   — **Feature Hypothesis:** The answer is a type of focus, which is marked with a morphosyntactic feature.

   — **Givenness Hypothesis:** The answer is prominent because it is not given—**given constituents resist prominence.**

In English, default prominence in neutral contexts falls on the rightmost constituent (11a), it is infelicitous for it to fall elsewhere (11b).

(11) **What happened?**
   a. My grandma made broccoli.
   b. #My grandma made broccoli.

However, constituents that are entailed by the context (i.e. **given**) resist phrasal prominence (Schwarzchild 1999, Féry and Samek-Lodovici 2006): Pitch accents cannot fall within the given constituent (12a), and instead fall on the new information (12b).

(12) **Who made the broccoli?**
   a. #My grandma [made the broccoli]\text{given}
   b. My grandma [made the broccoli]\text{given}

When the non-given constituent is a phrase, the pitch accent falls on the word with default phrasal prominence (13a) (Jackendoff 1972, Chomsky 1971).

(13) **Who made the broccoli?**
   a. My grandma from norway [made the broccoli]\text{given}
   b. #My grandma from Norway [made the broccoli]\text{given}

Crucially, the givenness hypothesis requires that the shifted prominence must map onto a position that is already prominent (Hayes 1995). The pitch accent must fall on the strong word within a phrase (14a) and the strong syllable within a word (15a).

(14) **What is your favorite food?**
   a. Raw broccoli [is my favorite food]\text{given}
   b. #Raw broccoli [is my favorite food]\text{given}

(15) **What is your favorite food?**
   a. broccoli [is my favorite food]\text{given}
   b. #Broccoli [is my favorite food]\text{given}
Superficially, a similar pattern appears to emerge in SMPM. When an answer to a wh-question is modified, preliminary evidence suggests that prominence falls on the adjective instead of the modified noun.

(16) Ná nákàbá nühü nü’ü
what fall.comp face ground
"What fell onto the ground?"

<table>
<thead>
<tr>
<th>a.</th>
<th>Tsáká [nákàbá nühü nü’ü]GIVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>fall.comp</td>
</tr>
<tr>
<td>&quot;The fish fell onto the ground.&quot;</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b.</th>
<th>Tsáká ndu’ù [nákàbá nühü nü’ü]GIVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>fat</td>
</tr>
<tr>
<td>&quot;The fat fish fell onto the ground.&quot;</td>
<td></td>
</tr>
</tbody>
</table>

Exs: tsyà ká (‘fish’) and tsyà ká ndu’ú (‘fat fish’)

Figure 12: High Tones in Fronted Position
Modified vs. Unmodified Nouns: n=242, p<0.05
Difference=0.14 ERB

On its face, this pattern is consistent with the givenness hypothesis, or with an analysis where the wh-answer is marked as a type a focus, and the prominence is realized at the right edge of the focus-marked phrase.

The givenness hypothesis, however, predicts that prominence must fall on the most prominent part of the non-given answer.

(17) Ná nákàbá nühü nü’ü
what fall.comp face ground
"What fell onto the floor?"

<table>
<thead>
<tr>
<th>a.</th>
<th>Tsyáká [nákàbá nühü nü’ü]GIVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>fall.comp</td>
</tr>
<tr>
<td>‘The FISH fell on the ground’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b.</th>
<th>Tsyáká ndu’ú [nákàbá nühü nü’ü]GIVEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>fat</td>
</tr>
<tr>
<td>‘The FAT FISH fell on the ground’</td>
<td></td>
</tr>
</tbody>
</table>
So, in order to account for the prominence pattern in (17), the **givenness hypothesis** predicts that in SMPM:

- Rightmost word is most prominent within the phrase, and;
- Rightmost syllable is most prominent within the word\(^5\)

(18)

\[
\varphi
\]

\[
\varphi_{\text{weak}}
\]

\[
\omega_{\text{strong}}
\]

\[
\sigma_{\text{weak}}
\]

\[
\text{tsyà}
\]

\[
\sigma_{\text{strong}}
\]

\[
ká
\]

\[
\sigma_{\text{weak}}
\]

\[
\text{ndu}
\]

\[
\sigma_{\text{strong}}
\]

\[
\text{'}ú'
\]

Crucially, *if this is a default prominence, then we expect it to also show up in non-focus contexts.*

- To test this prediction, I compared the pitch of H tones in modified nouns and adjectives in neutral contexts (Figure 13).

(19)  
Sháhmi (yútsí ihmí) rà lo'o  
burn sand hot he small  
‘The hot sand burned the boy.’

![Figure 13: Relative H tone Pitch of Nouns and Adjectives in Neutral Context](image)

\(n=128, \text{n.s.} \)

\(\text{Difference}=0.04 \text{ ERB} \)

---

\(^5\)Some support for this prediction comes from DiCanio et al. (2018), who argue for fixed final stress in Yoloxóchitl Mixtec. However, it should be noted that the towns of San Martín Peras and Yoloxóchitl are several hundred kilometers apart in a region with a great diversity of languages. More work needs to be done to determine if SMPM has a similar pattern.
There is **no evidence that adjectives are more prominent than the nouns they modify** in neutral contexts.

- Final high tones of adjectives are **not produced at a higher pitch** than final high tones of preceding nouns.

Alternatively, a proponent of the **givenness hypothesis** might say that there is an accent aligned to the right edge of intonational phrases in the language.

- When this position falls within a given constituent, the accent is aligned as close to the right edge as possible.
- This would result in the accent falling at the right edge of the non-given constituent.

If this is the case, **we should expect to see high tone prominence sentence finally in neutral contexts**.

- **Impressionistically, this is not the case**—there is declination throughout the clause and word-final high tones are produced at a lower pitch than high tones near the beginning of sentences.
- Additionally, **no speaker intuition that final word in a sentence is most prominent**.

I take these facts as **evidence against the givenness hypothesis**.

Instead, I suggest that **prosodic prominence is due to a feature, which triggers high tone raising at the right edge of the focused constituent**.

### 5 Syntactic Differences

Convergent evidence for the claim that wh-question answers are represented with a feature in the grammar comes from their syntactic behavior.

Recall that constituents can front when they are answers to wh-questions (20), or when they are part of the response to a broad focus question (21).

\[(20)\]
\[
\begin{align*}
\text{a. } & \text{Ná tāshī Mariana ntā’ā Bernardo} \quad \text{what give.comp M. hand B.} \\
& '\text{What did Mariana give Bernardo?}' \\
\text{b. } & \boxed{\text{Chichi}} tāshī=ňá ntā’ā=rā \quad \text{avocado give.comp=3sg.f hand=3sg.m} \\
& '\text{She gave him a \textit{avocado}.}'
\end{align*}
\]

\[(21)\]
\[
\begin{align*}
\text{a. } & \text{Ná kū bitsi} \quad \text{what cop.comp today} \\
& '\text{What happened today?}' \\
\text{b. } & \boxed{\text{Chichi}} tāshī Mariana ntā’ā Bernardo \quad \text{avocado give.comp M. hand B.} \\
& '\text{Mariana gave Bernardo an avocado.}'
\end{align*}
\]

The two hypotheses we are considering make different predictions with respect to the syntactic behavior of answers to wh-questions:

**Feature Hypothesis** predicts:
- Answers to wh-questions **should be subject to morphosyntactic operations in some languages**.

**Givenness Hypothesis** predicts:
- Answers to wh-questions are not formally represented in the grammar, so they **shouldn’t be involved in any special morphosyntactic operations**.

Independent syntactic evidence that the grammar is sensitive to answers to wh-questions comes from two diagnostics: optionality and clause-boundedness.
As we have seen, answers to wh-questions surface in a pre-verbal position. **This movement is obligatory.**

— When the object is questioned, it is ungrammatical to leave the answer in-situ (22b), to front the subject (22c), or to front a subject clitic pronoun (22d).

(22)  
\[ \text{Ná shishi tsiná} \]
what ate dog

‘What did the dog eat?’

a.  
\[ \text{Kōñù shishi=rí} \]
meat ate=it.animal

‘It ate the MEAT.’

b.  
\[ \text{*Shishi=rí kōñù} \]
ate=it.animal meat

Intended: It ate the MEAT.

c.  
\[ \text{*Tsiná shishi kōñù} \]
dog ate meat

Intended: The dog ate the MEAT.

d.  
\[ \text{*Rí=shishi kōñù} \]
it.animal=ate meat

Intended: It ate the MEAT.

This is not the case in broad focus contexts. Here, **either the subject (23a) or the object (23b) may front.**

(23)  
\[ \text{Ná kù bitsi} \]
what COP.COMP today

‘What happened today?’

a.  
\[ \text{Tsiná shishi kōñù} \]
dog ate meat

‘The dog ate the meat.’

b.  
\[ \text{Kōñù shishi tsiná} \]
meat ate dog

‘The dog ate the meat.’

This movement pattern strongly suggests that **the syntax must be able to uniquely identify answers to wh-questions** and that **there is a feature that must be checked.**

Second, fronting the answer to a wh-question is **not clause-bound.**

(24)  
\[ \text{Ná kàchi Juan [shàshi Ana]} \]
what said J. ate A.

‘What did Juan said that Ana ate?’

b.  
\[ \text{Ita, kàchi=rà [shàshi=ñá ____]} \]
flower said=3.sgm ate=3.sgm

‘He said that she ate a FLOWER.’

(25)  
\[ \text{Ná kàchi Juan [shí Ana nūhù yàbi]} \]
what said J. bought A. face market

‘What did Juan said that Ana bought at the market?’

b.  
\[ \text{Chíchí, kàchi=rà [shi=ñá ____]} \]
avocado said=3.sgm bought=3.sgf

‘He said she bought an AVOCADO’

Regardless of whether the answer originates in the matrix or embedded clause, it must move a sentence initial position.

6Interestingly, (1) is infelicitous in response to (23), but is acceptable as an out-of-the-blue assertion.

(1)  
\[ \text{#Shishi tsiná kōñù} \]
ate dog meat

Intended: The dog ate the meat.

I leave this fact as a puzzle for future research.
Fronting in a broad focus context, on the other hand, cannot proceed from the embedded clause.

(26) a. Ná kù bitsi
   what COP.COMP today
   ‘What happened today?’

b. *Ita kàchi Juan [shàshi Ana __]cp
   flower said J. ate A.
   ‘Juan said that Ana ate a flower.’

(27) a. Ná kù bitsi
   what COP.COMP today
   ‘What happened today?’

b. *Chichi kàchi Juan [shī Ana __]cp
   avocado said J. bought A.
   ‘Juan said that Ana bought an avocado.’

This syntactic data strongly suggests that answers to wh-questions are a category that the syntax cares about.

– Provides convergent evidence supporting the claim that answers to wh-questions are represented formally in the grammar.

– Similar to patterns of wh-movement in English (Chomsky 1977), and contrastive focus movement (E. Kiss 1998).

6 Conclusion and Future Directions

Today I have made the following claim: answers to wh-questions must be represented formally in the grammar of San Martín Peras Mixtec.

– This feature is realized by the phonology as asymmetric prominence.

– This feature triggers syntactic movement to a sentence initial position, even if it originates in an embedded clause.

Furthermore, I have argued against the givenness hypothesis on the following grounds:

– There is no evidence of default prominence in the language—a crucial component of that hypothesis.

– It incorrectly predicts that answers to wh-questions should not have special syntactic status.

Going forward, more work must be done to fit SMPM into the typology of prominence realization in tonal languages. Some tonal languages:

– Expand the entire F₀ space (e.g. Mandarin—Xu 1999).

– Shift the entire pitch register (e.g. Akan—Kügler and Genzel 2013).

– Realize prominence asymmetrically (e.g. SMPM, Hausa—Leben et al. 1989)

The asymmetric character of prominence in SMPM may help to adjudicate between differing theories of how answers to wh-questions receive their prominence cross-linguistically:

– Directly via spell-out at PF (Selkirk 1995, Rooth 1992, Féry and Samek-Lodovici 2006);

– Indirectly via alignment to a prosodic boundary (Büring 2013, Féry 2013, a.o.).
References


14


