Speech rate in phonology: Word reduction in San Martín Peras Mixtec

Ben Eischens, UCSC
SSILA 2021
Introduction

- Proposed “dividing lines” between phonology and phonetics:

<table>
<thead>
<tr>
<th>Phonology</th>
<th>Phonetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-specific</td>
<td>Universal</td>
</tr>
<tr>
<td>Contrastive</td>
<td>Non-contrastive</td>
</tr>
<tr>
<td>Abstract</td>
<td>Physical</td>
</tr>
<tr>
<td>Categorical</td>
<td>Gradient</td>
</tr>
</tbody>
</table>
Introduction

- Proposed “dividing lines” between phonology and phonetics.

<table>
<thead>
<tr>
<th>Phonology</th>
<th>Phonetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-specific</td>
<td>Universal</td>
</tr>
<tr>
<td>Contrastive</td>
<td>Non-contrastive</td>
</tr>
<tr>
<td>Abstract</td>
<td>Physical</td>
</tr>
<tr>
<td>Categorical</td>
<td>Gradient</td>
</tr>
</tbody>
</table>
Introduction

- Word reduction in San Martín Peras Mixtec (SMPM): Words with a laryngeal gesture between two homorganic vowels optionally reduce.

1. ŋini rà loʔo jìʔǐ koni
   Saw he small mush. yest.
   ‘The boy saw a mushroom yesterday’

2. ŋini rà lo jǐ koni
   Saw he small mush. yest.
   ‘The boy saw a mushroom yesterday.’
Introduction

● Word reduction in SMPM:
  ○ Gradient
  ○ Speech-rate dependent

● But…
  ○ Results in tonal restructuring
  ○ Feeds categorical phonology

● Phonetics interacting transparently with phonology
Overview

- Background
- The characteristics of word reduction
- Tone and tone sandhi
- Conclusion
Background
Background

San Martín Peras Mixtec (SMPM) is an Otomanguean language of Western Oaxaca, Mexico.

Spoken by ~10,000 in Oaxaca (INEG 2010)

~350,000 indigenous Oaxacans in California (Rabadán & Salgado 2018)
Ahuejutla, Oaxaca. Photo by Andrew Hedding
Background

Default VSO word order.

Tones: High (v́), mid (v), low (v̀), low-high rise (v̌), high-low fall (v̂).
- TBU is the mora

Tone is used in grammatical as well as lexical distinctions.
Background

Words in SMPM are minimally bi-moraic; no coda consonants.

Laryngeals are exceptional:
- The only “codas.”
- Culminative.
- Only occur root-medially.
Background

Non-laryngealized and laryngealized root shapes:

<table>
<thead>
<tr>
<th></th>
<th>CVCV</th>
<th>CVV</th>
<th>VCV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modal</strong></td>
<td>léló</td>
<td>lěè</td>
<td>ini</td>
<td>ĭī</td>
</tr>
<tr>
<td></td>
<td>“Skunk”</td>
<td>“Baby”</td>
<td>“inside”</td>
<td>“One”</td>
</tr>
<tr>
<td><strong>Laryngealized</strong></td>
<td>βîʔn̥tsià</td>
<td>kwîʔi</td>
<td>iʔnî</td>
<td>îʔî</td>
</tr>
<tr>
<td></td>
<td>“Nopal”</td>
<td>“Fruit”</td>
<td>“Hot”</td>
<td>“Raw”</td>
</tr>
</tbody>
</table>

Laryngealization in Mixtec is a supra-segmental feature, not a consonant (Macaulay & Salmons 1995, Gerfen 2013).
Background

Non-laryngealized and laryngealized root shapes:

<table>
<thead>
<tr>
<th></th>
<th>CVCV</th>
<th>CVV</th>
<th>VCV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modal</strong></td>
<td>léló</td>
<td>lěè</td>
<td>ini</td>
<td>ìî</td>
</tr>
<tr>
<td></td>
<td>“Skunk”</td>
<td>“Baby”</td>
<td>“inside”</td>
<td>“One”</td>
</tr>
<tr>
<td><strong>Laryngealized</strong></td>
<td>βìʔⁿtsià</td>
<td>kwìʔi</td>
<td>iʔní</td>
<td>ìʔi</td>
</tr>
<tr>
<td></td>
<td>“Nopal”</td>
<td>“Fruit”</td>
<td>“Hot”</td>
<td>“Raw”</td>
</tr>
</tbody>
</table>

Laryngealization in Mixtec is a supra-segmental feature, not a consonant (Macaulay & Salmons 1995, Gerfen 2013).
Background

Phonological representation of CVʔV:

1. tsjáʔa
   “Tecomaxtlanahuaca”
Background

Methods:

- Data collection:
  - Most data come from one speaker with auxiliary data from another.
  - Audio collected online using Zencastr recording software and a Cooler Master MH630 microphone.
  - Tasks included translations, well-formedness judgments, informal forced-choice tasks, etc.
  - Aggregated pitch data comes from one production task
    - Five repetitions of target word in carrier sentence, followed by five repetitions of target word in reduced form in same sentence.
Characteristics of word reduction
Word reduction: Distribution

Words with a laryngeal gesture between two homorganic vowels optionally reduce.

1. ŋĩi rà loʔo ŋiʔĩ koni
   Saw he small mush. yest.
   ‘The boy saw a mushroom yesterday’

2. ŋĩi rà lo ŋi koni
   Saw he small mush. yest.
   ‘The boy saw a mushroom yesterday.’
Word reduction: Distribution

This process applies to words of different syntactic categories.

<table>
<thead>
<tr>
<th>Verbs</th>
<th>Prepositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>nũhnĨ tʃiʔi  rã Corn plants he ‘He’s planting corn’</td>
<td>tàʃĨ Ped. jùʔũ ndàʔã Franc. gave P money hand F ‘Pedro gave money to Francisco’</td>
</tr>
<tr>
<td>nũhnĨ tʃi  rã Corn plants he ‘He’s planting corn’</td>
<td>Gave P money hand F ‘Pedro gave money to Francisco’</td>
</tr>
</tbody>
</table>
Word reduction: Distribution

It happens less often in prosodically strong positions like…

- Utterance-final

- Focus-fronted

1. tàʔβi tjútu kòʔǒ~#kǒ
   Broke cat plate
   ‘The cat broke a plate.’

2. βeʔe~#βe nàkaba nũ 琎ũʔũ
   House fell face earth
   ‘The house fell to the ground’
But even in these positions, it is possible, especially if the word is previously mentioned.

Speaker 1: tàʔβi tjútu kòʔǒ?
Broke cat plate
‘Did the cat break a plate?’

Speaker 2: āhā, tàʔβi tjútu kǒ
yes, broke cat plate
‘Yes, the cat broke a plate.’
Word reduction: Distribution

In fact, reduced words can be used as fragment answers.

Speaker 1:  kʷìʔi  ra ḟiʔi  βa  nì-ʃìhjó  nũhũn  maría
Fruit and mush. EMPH existed face Maria
‘María had fruit and a mushroom.’

Speaker 2: ⁿtsiàà  já ḟàfi  ŋá?
Which it ate she
‘Which did she eat?’

Speaker 1:  kʷi
Fruit
‘Fruit’
Word reduction: Phonetic characteristics

Word reduction occurs most often in fast speech.

- Informal production task varying prompted speech rate and previous mentions

<table>
<thead>
<tr>
<th></th>
<th>Fast production</th>
<th>Slow production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previously mentioned</td>
<td>18/36</td>
<td>0/34</td>
</tr>
<tr>
<td>Not previously mentioned</td>
<td>12/34</td>
<td>2/36</td>
</tr>
</tbody>
</table>

Rates of reduction by speech rate and givenness
Word reduction: Phonetic characteristics

Word reduction is also gradient: There is a cline of laryngealization from glottal closure to apparent modal voice:

Four productions of the vocalic portion of the word *loʔo* (“small”) in the same syntactic position by the same speaker.
Word reduction: Phonetic characteristics

Review:

- Word reduction can apply to words of all syntactic categories in all prosodic positions.
- Word reduction has phonetic characteristics:
  - Gradient
Tone and tone sandhi
Tone and tone sandhi

Word reduction in SMPM comes along with tonal changes.

Tones usually appear to merge.

1. tsjàʔá ~ tsjà
   Salsa
   `Salsa'
Tone and tone sandhi

Word reduction in SMPM comes along with tonal changes.

Tones usually appear to merge.

1. tsiàʔá ~ tsiã
   Salsa  `Salsa`

Pitch (Hz) for reduced and unreduced productions of laryngealized words with a L-H melody. n=61 (31 Long, 30 Short).
Tone and tone sandhi

Tone sandhi: A word-final LH contour flattens to L when followed by a H:

1. Ni-ⁿtsikù tsjùhũ leso
   Chased turkey rabbit
   ‘The turkey chased the rabbit.’

2. Ni-ⁿtsikù tsjùhũ léló
   Chased turkey skunk
   ‘The turkey chased the skunk’
Tone and tone sandhi

tsiûhû lélé → tsìûhû lélé

This is phonological:
- Categorical at all speech rates
- Neutralizes between LH and L
- Does not apply to negative grammatical LH tone
Tone and tone sandhi

This only applies to LH contours, not sequences of L and H.

1. Ḗni tsǐĩ leso
   Saw rat rabbit
   “The rat saw the rabbit”

2. Ḗni tsǐĩ léló
   Saw rat skunk
   “The rat saw the skunk”
Tone and tone sandhi

\( \text{tsǐ̝́́ léló} \nswap \text{tsǐ̝́́ léló} \)
Tone and tone sandhi

Reducing L-H words creates an apparent rise:

1. shǐni rà loʔo tsiòʔó koni
   Saw he small flea yest.
   ‘The boy saw a flea yesterday.’

2. shǐni rà lo tsiõ koni
   saw he small flea yest.
   ‘The boy saw a flea yesterday.’
Tone and tone sandhi

That derived rise undergoes tone sandhi.

1. shîni rà loʔo tsiòʔó tákáʔe
   Sees he small flea now
   ‘The boy sees a small flea now.’

2. shîni rà loʔo tsiò tákáʔe
   Sees he small flea now
   ‘The boy sees a small flea now.’
Tone and tone sandhi

This change is relatively consistent.

Pitch (Hz) before an H tone for reduced and unreduced productions of laryngealized words with a L-H melody as well as vowels with an underlying L tone. n=41 (14 Long, 13 Short, 14 Low)
Tone and tone sandhi

Interim review:

Sandhi does not apply to unreduced laryngealized L-H words.
- Like in bi-moraic CVV words.

Sandhi does apply to reduced laryngealized L-H words.
- Like in mono-moraic syllables with a LH contour.

This implies a difference in tonal alignment.
Tone and tone sandhi

Unreduced laryngealized words have a similar structure to CVV words:

Because these are sequences of L and H, they do not undergo sandhi.
Tone and tone sandhi

Tone sandhi applies to mono-moraic LH rises:

Since tone sandhi applies to reduced laryngealized words, they must be mono-moraic.
Tone and tone sandhi

A mono-moraic structure suggests that word reduction looks something like this:
Tone and tone sandhi

A mono-moraic structure suggests that word reduction looks something like this:
Tone and tone sandhi

Word reduction involves a change to the tonal representation.
  - Mora deletion, followed by tonal reassociation.
  - Tonal reassociation feeds phonological tone sandhi.

Conclusion: **Word reduction is phonological.**
  - Despite its gradience and sensitivity to speech rate.
Tone and tone sandhi

Review:
- Word reduction has the characteristics of a phonetic process:
  - Gradient.
  - Conditioned by speech rate.
- It also involves mora deletion and tonal reassociation, suggesting it is phonological.
  - Feeds phonological tone sandhi.
Word reduction is gradient and determined by speech rate (physical), but it can result in categorical change to the abstract tonal representation.
Conclusion

If a modular view to the phonology-phonetics interface is still to be upheld, they must interact in some way.

- There must be some level in the synchronic grammar where phonetic factors like speech rate can influence abstract phonological computation.
Conclusion

Another takeaway:

- Being a “fast-speech” process does not entail being a “phonetic” process (contra McCarthy 1986; Keating 1996; Myers 2000).
- A gradient process can correspond to categorical phonological change (McCollum, 2019, Du and Durvasula 2020).
Thank you!

Thank you to Eraclio Gracida Cruz, Juan Gracida Ortiz, Irma López Basurto, one anonymous consultant, Roselia Durán Cruz, and most especially Natalia Gracida Cruz for their patience and enthusiasm in sharing their language with me.

Project funded in part by a Summer Research Fellowship from The Humanities Institute at the University of California, Santa Cruz.
Appendix
A. Tone and tone sandhi

In most other cases, tones appear to merge

Pitch (Hz) for reduced and unreduced productions of one laryngealized word with a H-L melody. n=31 (24 Long, 17 Short).

Pitch (Hz) for reduced and unreduced productions of laryngealized words with a H-M melody. n=36 (23 Long, 23 Short).
Tone and tone sandhi

Tones often appear to merge:

Pitch (Hz) for shortened and unshortened productions of laryngealized words with a M-LH melody. n=90 (46 Long, 44 Short).

Pitch (Hz) for shortened and unshortened productions of laryngealized words with a L-LH melody. n=46 (22 Long, 23 Short).
Tone and tone sandhi

In one tone class, it looks like a tone deletes.

1.  tùʔũ word  `Word’
Tone and tone sandhi

In one tone class, it looks like a tone deletes.

1. ตุ้ง.th
   word  ‘Word’
Tone and tone sandhi

Process 2: A word-initial high becomes a rise when following a low

1. jéʔe káʔnũ
door big
‘A big door.’

2. máʔà káʔnũ
racc. Big
‘A big raccoon.’
Tone and tone sandhi

When reduction derives this environment, sandhi does not happen.

1. shën-ì  tsiòʔo  káʔnũ
   saw-l  root  big
   ‘I saw a big root.’

2. shën-ì  tsiò  káʔnũ
   Saw-l  root  big
   ‘I saw a big root.’
Tone and tone sandhi

Reduction feeds process 1 but counterfeeds process 2.

<table>
<thead>
<tr>
<th>Process 1</th>
<th>Process 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully general</td>
<td>Lexically-specified</td>
</tr>
<tr>
<td>Post-lexical</td>
<td>Has an analogue in lexical phonology</td>
</tr>
<tr>
<td>Applies across larger prosodic boundaries</td>
<td>Applies across fewer prosodic boundaries</td>
</tr>
<tr>
<td>than Process 2</td>
<td>than Process 1.</td>
</tr>
</tbody>
</table>
Tone and tone sandhi

It is entirely possible that in reduced L-M words, the fast rate of speech results in undershoot of the M tone, analogous to what happens in LML CVV words.
B. Laryngealization

Reduced words lack obvious laryngealization:

Intensity (normalized) across the duration (normalized) of the root for shortened and non-shortened roots with a laryngealized vowel as well as for modal CVV roots. n=about 500.

Vocalic portion of long and short forms of the word loʔo (‘small’)

Intensity contours by word type
Laryngealization

Reduced forms also show less f0 modulation.

Pitch (Hz) for shortened and unshortened productions of laryngealized words with a H-M melody. n=36 (23 Long, 23 Short).
Laryngealization

They still appear to be laryngealized:

H1-H2 (dB, uncorrected) for the vocalic portions of unshortened laryngealized words and modal, CVCV words, alongside H1-H2 for shortened laryngealized words. n= 32 each.
Laryngealization

Other measures are not so clear:

CPP for the vocalic portions of unshortened laryngealized words and modal, CVCV words, alongside H1-H2 for shortened laryngealized words. n= 32 each.
Laryngealization

Other measures are not so clear:

CPP for the vocalic portions of unshortened laryngealized words and modal, CVCV words, alongside H1-H2 for shortened laryngealized words. n= 32 each.
Laryngealization

Review:

- Reduced words lack creak, amplitude dips, and f0 modulation.
- H1-H2 shows reduced words patterning with unreduced laryngealized words; CPP and HNR inconclusive.
C. Variable application of tone sandhi

There are some syntactic environments where tone sandhi applies to underlying LH rises but not LH rises derived by word reduction:

1. ʃǐni rà loʔo kòhò kʷíhi koni
   see.Compl he small snake green green yesterday
   “The boy saw a green snake yesterday”

2. ʃǐni rà loʔo tsʲòˀó/tsʲǒ kʷíhi koni
   see.Compl he small flea green green yesterday
   “The boy saw a green flea yesterday”
Variable application of tone sandhi

Pitch (Hz) for reduced and unreduced productions of laryngealized words with a L-H melody, as well as vowels with an underlying LH contour or L tone, before an H-initial adjective. N=90 (19 Rise, 24 Low, 23 Long, 24 Short).
Variable application of tone sandhi

In other syntactic environments, application of tone sandhi to derived rises is less consistent than to underlying rises.

Pitch (Hz) for reduced and unreduced productions of laryngealized words with a L-H melody before an H-initial adverb. N=37 (18 Long, 19 Short).
Variable application of tone sandhi

The environment in which sandhi seems to always apply to derived rises is in Possessive structures:

1. kón=ì tsjàʔá léló want.Cont=1Sg salsa skunk “I want the skunk’s salsa”

2. kón=ì tsjà léló want.Cont=1Sg salsa skunk “I want the skunk’s salsa”