Challenging WEIRD phonetic assumptions: phonetic fieldwork in rural Mexico

Christian DiCanio
cdicanio@buffalo.edu

Department of Linguistics
University at Buffalo

1/7/22 and 2/10/22
Methods sometimes fail

Typical assumptions and constraints in experimental design may either be inapplicable or impossible to implement in field contexts.

The field strongly benefits from phonetic fieldwork, but our challenge is maintaining careful design standards for in-situ field sites.
Outline

1. WEIRD field assumptions and solutions
   a. Case study (literacy): prosodic boundary elicitation while controlling for tonal context
   b. Case study (imitation): delayed shadowing tasks and methods to avoid imitation

2. Solutions for design difficulties
   a. Literacy and language learning goals which overlap with experimental or exploratory goals
   b. Flexibility in preregistration
Languages spoken in WEIRD societies (Western, educated, industrialized, rich and democratic) (Henrich et al., 2010) dominate research in phonetics.

Studies on (majority dialects of) Western Indo-European languages continue to dominate phonetic research.

The development of more advanced experimental methods often relies on a set of assumptions true for these languages (but less true for others), especially as related to educational background.
Methodological assumptions and phonetic fieldwork

JIPA Illustrations from 1989 - 2018
(Whalen, DiCanio, and Dockum 2020)

Of the approximately 1,560 articles published, 110 were deemed to be potentially relevant to our study on the basis of the title and abstract. These were examined in more detail by two of the authors for a judgment of whether the article could count as documentation, and initial disagreements were discussed until a decision was made. Measurements of at least an entire category in the phonemic inventory were necessary at a minimum, e.g. the vowels or nasal consonants of the language. The amount of material and coverage needed was not fixed, and the decisions necessarily rested on multiple considerations. The number of speakers and tokens measured was one feature, with larger numbers, naturally, making it more likely that a 'yes' judgment would be made. Measurements of a single aspect of a distinction (such as perturbations of fundamental frequency (f0) by stop voicing) were generally insufficient, while more global (and common) measures such as voice onset time (VOT) would be sufficient. It is always possible to examine the interactions between different phonetic distinctions, but if such interactions were the primary goal of the article, it was excluded from consideration. For instance, a description would ideally contain measurements for all members of a particular phonological category, e.g. all stops, all nasals, or all coronal consonants. Those articles specializing on a small subset of the phonetic contrasts in a series were excluded. The presence or absence of a theoretical conclusion was not taken into account; most of the articles did indeed make a theoretical point. Articles that dealt exclusively with perception were excluded; see the 'Discussion' section for further considerations about perception.

Coverage for the categories ranges from 2.2% to 39.1%, with a mean of 12.5% and a standard deviation of 7.4%; see Table 3 for a breakdown by category. VOT (29.7%) and both closure duration and voicing (each 20.3%) were the most commonly reported consonantal categories; 'vowel features' (25.0%) and formants (22.7%) for the vowels; and tone/pitch accent (17.2%) and interactions (12.5%) for the suprasegmentals. There was an uptick in coverage between the 1990s (7.4%) and the first decade of the 2000s (16.1%), but then a slight decline in the 2010s (13.1%). As with JIPA, there is a variation in each of those ranges.
WEIRD problems

- Most work on the phonetics of prosody involves a reading task with careful control of prosodic context, but only $\sim10\%$ of the world’s languages have literacy (Harrison, 2007).

- Tasks involving goodness ratings or evaluation of non-linguistic stimuli may fail with populations unfamiliar with experiments in an educational setting.

- Close attention to phonetic detail in perception is influenced by literacy (McGuire, 2010). This will predict different outcomes in listening and shadowing tasks with illiterate speakers than with literate speakers.
Case studies on the phonetics of prosody in Mixtecan languages

- Research on the prosodic structure of Itunyoso Triqui (ISO trq) and Yoloxóchitl Mixtec (ISO xty) in rural Mexico.

- Both languages are Otomanguean: Mixtecan and have approximately 2,500 speakers; mostly illiterate in their native language. These are heavily tonal (9+ tones) languages.

- Higher-level prosodic contrasts (information structural, positional, intonational) were investigated in the field from 2015 - 2019.

Challenge: Controlling for tone, phonation type, word length, and context is necessary when examining prosody.
Several studies examining prosodic focus involve reading tasks e.g. in Mandarin (Chen and Gussenhoven, 2008; Xu, 1999), Guaraní (Clopper and Tonhauser, 2013), Arabic (de Jong and Zawaydeh, 2002), German (Mücke and Grice, 2014), Dutch (Peters et al., 2014)

An alternative with illiterate speakers of Uspanteko (Bennett et al., 2021), Itunyoso Triqui (DiCanio and Hatcher, 2018), Yoloxóchitl Mixtec (DiCanio et al., 2018) - short narrative texts followed by question-and-answers.

Speakers successfully answer questions related to the text without many problems.
Methodological assumptions and phonetic fieldwork

Triqui example in tonal orthography

4. Ngo2 chu3ba4 ka3hanj3 na2noh2=chuj3 ni3ya32 cha2=chuj3. Ki3ni3hin3=chuj3 kkan3 ri3ki3 nna31 ngaj23 na4taj4. Yya3 ka3chin3 chi3hna32 ri3ki3=chuj3 ni2 chu3ba4 cha43 ba1hni3. Ni2 na3bin3 ya3hmanh3=chuj3 ni2 ki3ni3hin3 cha3ta32 si3 ba2 sah1=chuj3 ni2 cha43 cha3ta32=chuj3.

Una ardilla fue a buscar comida. Vio calabazas en la milpa que estaba en la loma. Desde que tuvo mucha hambre, la ardilla comió tres. Por ponerse muy gordo la águila lo vio mejor y se lo comió.

a. Un3 chu3ku3 ka3hanj3 na2noh2 ni3ya32 cha2 oh2?
   Qué fue el animal que fue a buscar comida?

b. Un3 sin3 cha43 chu3ba4 oh2?
   Qué comió la ardilla?

c. Un3 cchej32 ngaj23 nna31 rian32 na3rih3=chuj3 kkan3 oh2?
   Dónde estaba la milpa donde encontró las calabazas?

d. Un3 taj3 kkan3 cha43 chu3ba4 oh2?
   Cuántas calabazas comió la ardilla?

e. Un3 chu3ku3 cha43 chu3ba4 oh2?
   Qué animal comió la ardilla?

Blancos:
chu3ba4    'ardilla'
na4taj4    'loma'
ba1hni3    'tres de algo'
kkan3      'calabaza'
cha3ta32   'águila'

Translation

Translation

Questions

Intended
targets from
text
Still requires some language literacy.

Either a speaker is trained in literacy and reads the text and questions aloud –or– they record a short narrative in the language and translate questions into the target language for other speakers.

**Potential issue 1:** Speakers may produce single word responses rather as “complete sentences.” Researchers need to ask speakers to produce “full sentences” with a highlit NP, but this may be less natural.

**Potential issue 2:** This method works well for eliciting contrastive or narrow focus, but fails with broad focus.
Elicitation of prosodic focus in the field

- Narrow focus (after story)
  Rey: Who arrived?
  Speaker (instructed): John arrived.
  Speaker (more natural): John

- Contrastive focus (after story)
  Rey: Did Marcus arrive?
  Speaker (instructed): John arrived.
  Speaker (more natural): No, John did.

* Sentential/broad focus
  Consultant: What happened?
  Speaker: You just told me what happened. Why are you asking me?
In field contexts where broad focus was elicited, both Triqui speakers and Mixtec speakers failed to understand the task’s motivation.

Either no response was given or speakers chose to provide what they saw as the “moral of the story.” A question requiring dictation of information that was just given probably does not arise outside of contexts involving literacy.

Potential solution 1: Use of existing corpora to find words under broad focus. Though, if you have to match words for tone/length, this option usually is impossible.

Potential solution 2: A repetition task for broad focus.
Repetition/shadowing tasks in the field

The easiest solution? A speaker simply repeats the target sentence under broad focus.

**Potential issue:** Speakers of English and Italian may imitate prosodic contrasts in human language (Cole and Shattuck-Hufnagel, 2011; D’Imperio et al., 2014).

Whether speakers of (a) tonal languages or (b) minority illiterate populations imitate prosodic detail in repetition tasks remain open questions in phonetics.
Potential solution: delayed shadowing

Sensory memory traces containing detailed phonetic information decay within 300 ms and analyzed sensory memory traces to decay over a slightly longer timecourse (Pisoni, 1973; Xu et al., 2006).

A potential solution here is for speakers to produce a fixed phonetic sequence prior to repeating a target sentence.

e.g. “One, two, three.... Maria arrived today at 2 PM.”

Goal: the participant reproduces the lexical content (allowing for careful phonetic design) but not the phonetic details from the speaker.

(Acknowledgement: Lisa Davidson)
Methodological assumptions and phonetic fieldwork

Speaker/listener imitation

Speakers hesitate, but respond well

Speech sample

[Diagram showing a speech sample with a stimulus and responses]
Recruiting a confederate (h/t Katie Franich)

Another alternative that has not been attempted is recruitment of a non-participating confederate in the dialogue between native speakers (Hwang et al., 2015).

Design: a non-participant could come into the room and ask the participant to re-tell the narrative that they heard.

Potential problems:

- To what extent would participants remember all the relevant details (repeat the target words)?
- It is awkward to ask a person to leave and come back for each set of questions (to ensure they do not hear the story).
What about mining corpora for examples?

In the context of speech prosody, it is often necessary to carefully control for (a) rhythmic and accentual structure (cf. Turk and White, 1999) and (b) word-prosodic features like tone.

While corpus-based phonetic methods have advanced, it is quite difficult to find a substantial set of tonal pairs in different prosodic positions in the same recording.

(Ignoring that you need a transcribed and aligned corpus - possibly substantial.)
Case study: Mixtec phrase-final lengthening

In a recent study, we examined the interaction between tone and utterance positions (medial/final) in Yoloxóchitl Mixtec. Using a repetition task, we could examine 20 distinct tonal melodies (DiCanio et al., 2021).

| Table 1: Experimental stimuli - tonal melodies in disyllabic words. Note that tones are specified on individual syllables here, not words, for instance, \( /\beta a^{14}\text{ni}^{4}\) “rainbow” has an initial rising tone followed by a high tone. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Melody | Word | Gloss | Melody | Word | Gloss |
| 4.4 | \( /\beta a^{4}\text{li}^{4}\) | “little” | 3.4 | \( /\text{di}^{3}\text{ji}^{4}\) | “corncob” |
| 4.3 | \( /\text{ju}^{4}\text{nu}^{4}\) | “pineapple” | 3.3 | \( /\text{t}^{3}\text{ji}^{3}\text{u}^{3}\) | “work” |
| 4.3 | \( /\beta i^{4}\text{lu}^{3}\) | “clay” | 4.2 | \( /\text{ta}^{3}\text{ta}^{2}\) | “banana” |
| 4.2 | \( /\text{di}^{4}\text{ko}^{3}\) | “mother” | 4.1 | \( /\text{ta}^{3}\text{ta}^{2}\) | “seed” |
| 4.2 | \( /\text{t}^{3}\text{ji}^{4}\text{tu}^{2}\) | “woodpecker” | 4.1 | \( /\text{ta}^{3}\text{ta}^{2}\) | “seed” |
| 4.1 | \( /\beta i^{4}\text{f}^{i}^{1}\) | “difficult” | 13.3 | \( /\text{ki}^{13}\text{f}^{i}^{3}\) | “come.PERF” |
| 4.1 | \( /\text{ta}^{4}\text{si}^{1}\) | “wizard” | 13.2 | \( /\text{ju}^{13}\text{du}^{2}\) | “doll” |
| 14.3 | \( /\text{t}^{4}\text{f}^{i}^{4}\text{ku}^{3}\) | “guamuchiles” | 14.3 | \( /\text{t}^{4}\text{f}^{i}^{4}\text{ki}^{3}\) | “prickly pear” |
| 14.2 | \( /\text{ja}^{4}\text{ku}^{2}\) | “be itchy” | 4.13 | \( /\text{t}^{4}\text{f}^{i}^{4}\text{si}^{13}\) | “straight” |
| 1.42 | \( /\text{ta}^{4}\text{k}^{4}\text{wi}^{42}\) | “water” | 14.4 | \( /\beta a^{14}\text{gi}^{4}\) | “rainbow” |
But only three tonal melodies could be examined in a naturalistic corpus (the frog story). Matched utterance-final words were difficult to find.

<table>
<thead>
<tr>
<th>Table 3. Stimuli chosen from corpus.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melody</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>3.4</td>
</tr>
<tr>
<td>3.2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1.1</td>
</tr>
</tbody>
</table>
Similar effects of prosodic lengthening were found (results from repetition on left, the corpus results on right).

Note that the scales are slightly different (running speech is faster).
Summary

Phonetic fieldwork often involves work with illiterate populations.

Careful control of elicitation frames for specific research questions can be accomplished via naturalistic tasks (stories, Q&A) and, where needed, adjustments to simpler repetition tasks.

For more exploratory phonetic research requiring less control, corpus phonetic methods are useful.
Long-term collaborative work → better design

Not all phonetic studies are directly important to language communities.

Those which are especially useful examine naturalistic contexts encountered by the language community.

Collaborative goals are often clearer to speakers if they are coupled with literacy or pedagogical outcomes.

e.g. production/perception of contrasts that might be otherwise difficult to perceive; recordings that could be used as entries in a dictionary or language learning website.
The Itunyoso Triqui literacy website contains examples of tones in distinct contexts. These contexts serve to illustrate which tones are written and which ones are not.

Ejemplos de palabras con tonos diferentes (de nivel) están acá abajo:

- **anin**: 'estallar' (tono 1 - más bajo)
- **achin**: 'faltar' (tono 3 - medio)
- **achinj**: 'pedir' (tono 45 - lo más alto)
- **chihi**: 'enfermedad' (tono 1 - más bajo)
- **chi**: 'abuelo de' (tono 3 - medio)
- **chihi**: 'sobre' (tono 4 - alto)
Working with speakers

Your knowledge of the language permits experimental flexibility. Speakers have the most knowledge.

Literacy training is capacity-building for future pedagogy and cross-collaboration.

It is usually more natural for a speaker to read a text in the context of an experiment (and to explain the goals) than it is for a non-speaker to do so.

Training in literacy allows for better collaboration, but also necessitates a long-term investment in studying individual languages.
Preregistration in linguistic research

“A preregistration is a time-stamped document in which researchers specify prior to data collection how they plan to collect their data and/or how they plan to conduct the data analyses.” (Roettger, 2021, 1228)

The primary goal of preregistration is to track analytical flexibility and counteract publication bias in design.
Preregistration and design flexibility

Exploratory field phonetics does not require preregistration since it does not involve hypothesis-testing (Roettger, 2021).

Though, careful design is not limited to confirmatory research in field phonetics (prosody!).

Exploratory studies which involve careful design aspects should probably be preregistered. After all, aspects of elicitation design are not the primary reason for preregistration - the focus is more on statistical methods and post-processing.
Not all exploration is created equal.

**Exploration type 1:** initial stage exploration without much background or knowledge of patterns which interact with each other, e.g. the field linguist has identified phonological contrasts and wishes to record words with these contrasts in different positions of the word.

Probably **not** preregistered.

**Exploration type 2:** involves very intricate knowledge of many aspects of the language’s structure (morphosyntax, word-level contrasts), such as prosody/phrasing, intonation, and even certain socio-phonetic features.

Probably preregistered.
Too little flexibility might prevent the types of useful in-the-field adjustments illustrated in the prosodic studies mentioned here.

An ideal scenario – reviewers permit some flexibility in preregistration.

What phonetic fieldwork is preregistered? An evolving question.
Final remarks

Idealized phonetic research methods may involve biased assumptions.

Read speech itself reflects a speech style that is distinct from spontaneous speech - it is not “neutral” (cf. DiCanio et al., 2021).

Close collaborations with speakers (and literacy training) fosters alternative designs.

Preregistration is not inflexible.
Thank you!


Conclusions


