Speech data acquisition - The underestimated challenge

Tutorial at the Symposium on Tonal Aspects of Language
Nanjing, China, 26 May 2012

Oliver Niebuhr
Analysis of Spoken Language, University of Kiel, Germany
&
Alexis Michaud
CNRS – LACITO & CEFC / Université Sorbonne Nouvelle, France
Introduction

• Speech is recorded...
  • to document little-described languages  
  • to develop & test theories about specific forms and functions in a given language.

• “Field” and “lab” are nowadays more research concepts than research locations.
  • Data can be acquired and analyzed outside the lab, “in the field”
    Portable devices include: sound recorder; electroglottograph; electropalatograph; Ultrasound.
    Software: spectrograms, F0 extraction...
  • Speakers of a little-documented language can be brought to a lab (in some cases)
Introduction

• Data-oriented ↔ theory-oriented rather than “field” ↔ “lab”
• The factors that determine the richness, reliability and (ecological) validity of the data are similar.

• Aim of tutorial: to sensitize researchers to the problems and possibilities in the acquisition of speech data
  • Outlining ways in which experimenters in the lab and field workers can benefit from one another. (Alexis Michaud)
  • Providing an overview of tasks that can be used to guide elicitation. (Oliver Niebuhr)
Introduction

Organization of the tutorial:
- Part 1: data acquisition in the field
- Part 2: theory-oriented speech-data acquisition
  - → the 4 cornerstones of a speech corpus
  - → 1 way to exploit them
Part 1: data acquisition in the field
Part 1: data acquisition in the field

‘Classical’ fieldwork:
- vocabulary list,
- sentences,
- transcription of narratives.
Contents of the archives (How to / Search)

The LACITO Archive provides free access to documents of connected, spontaneous speech, mostly in "rare" or endangered languages, recorded in their cultural context and transcribed in ...

More

The public Archive contain over 200 documents in 45 languages, annotated by some twenty specialists (this list does not take into account the documents which are currently being processed)

- African languages
- Balkan languages
- Caucasus
- Organic languages

Nepal/Asia Languages

Corpus of African languages (introduction)

<table>
<thead>
<tr>
<th>Language</th>
<th>Family</th>
<th>Country</th>
<th>Researcher(s)</th>
<th>Archive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ihanzu</td>
<td>Bantu</td>
<td>Tanzania</td>
<td>Margaret Dunham</td>
<td></td>
</tr>
</tbody>
</table>
The Limbu inhabit Nepal east of the Arun river and bordering areas of the Indian states of West Bengal (Darjeeling District) and Sikkim. They are called "Limbu" in Nepali, "Yakthung" in Limbu. There are some 200,000 speakers of the language, virtually all bilingual in Nepali, the Indo-Aryan national language of Nepal.

The language

The Limbu language belongs to the Tibeto-Burman family. It was classified by Shafer (1955) in the East Himalayish section of the Bodic Division of Sino-Tibetan. Limbu is the easternmost member of this group, also known as "Kiranti". There are several Limbu dialects, but speakers generally consider them to be intercomprehensible, with a bit of effort.

The following works on Limbu may be found useful:

Once long back, in Linkhim in the Tamur Valley, when I was staying up at my father-in-law's place.

"Allons couper du bois à brûler" oint dit mon beau-père et mes beaux-frères.

With us were Sakcyong, whom I'd brought along [from Libang] and me, with Panther Mahila and Mudamba — we were staying there. Avec nous, il y avait Sakcyong, que j'avais amené [de Libang], moi, et aussi Panther Mahila et Mudamba - on était tous là.
Where the shoe pinches:
Part 1: data acquisition in the field

Should phonetic/phonological research be conducted independently?

Narratives may be inadequate
- for phonological purposes

Exploration of the tones of compound nouns in Yongning Na:
  requires 16x16 combinations.
  Same for noun+verb, etc.

- for phonetic purposes
  • Good audio
  • Airflow?
  • Electroglottography?
  • Articulatography?
  • Video?

> separate work?
Part 1: data acquisition in the field

'Classical' fieldwork:
- vocabulary list,
- sentences,
- transcription of narratives.

Oliver Niebuhr & Alexis Michaud
Part 1: data acquisition in the field
Part 1: data acquisition in the field

‘Practical wisdom’:

- Acoustic phonetics: central part of the discipline. Major difference: with or without sound (=with or without spectrograms).
- Quality recordings can be conducted in the field
- Abundant materials can be collected
- In-depth collaboration with language consultants is possible

An example: the morpho-tonology of numeral-plus-classifier phrases. (Language: Yongning Na)
<table>
<thead>
<tr>
<th>W8</th>
<th>W9</th>
<th>W10</th>
<th>W11</th>
</tr>
</thead>
<tbody>
<tr>
<td>hū₁-[m]</td>
<td>gyl-[mɪ]</td>
<td>tsʰe₁-[m]</td>
<td>tsʰe₁-[m]</td>
</tr>
<tr>
<td>8 粒 (米), 个 (碗), 件 (衣服)</td>
<td>9 粒 (米), 个 (碗), 件 (衣服)</td>
<td>10 粒 (米), 个 (碗), 件 (衣服)</td>
<td>11 粒 (米), 个 (碗), 件 (衣服)</td>
</tr>
<tr>
<td>8 + classifier for round objects; used as generic classifier</td>
<td>9 + classifier for round objects; used as generic classifier</td>
<td>10 + classifier for round objects; used as generic classifier</td>
<td>11 + classifier for round objects; used as generic classifier</td>
</tr>
<tr>
<td>The transcription is a surface phonological representation. The underlying tone of this phrase is: HS.</td>
<td>The transcription is a surface phonological representation. The underlying tone of this phrase is: L.</td>
<td>The transcription is a surface phonological representation. The underlying tone of this phrase is: M.</td>
<td>The transcription is a surface phonological representation. The underlying tone of this phrase is: M.</td>
</tr>
</tbody>
</table>
Part 1: data acquisition in the field

An ‘odd-man-out’ in the early stages of a description-oriented project.

But a useful part of in-depth fieldwork.

“Phonotactics and the prestopped velar lateral in Hiw: Resolving the ambiguity of a complex segment”, *Phonology* 27.3 (2010), by Alexandre François

State-of-the-art phonetic data can be collected during fieldwork, thereby improving the record.

Facilitates scientific communication.

Preserving the data & making them available: cumulative progress.
Conclusion of Part 1 (data acquisition in the field): advice

Short-term perspective: Practise recording.

Mid-term perspective: Go a few steps out of your way, for phonetics’ sake.

Long-term perspective: Put the lot online.
Part 2 « Theory-oriented Speech-Data Acquisition »

- Question / Aim
  - documentation
  - investigation

- Subject
  - (grammar, phonology, sounds, prosodies)

- Informants/speakers
- Speech data acquisition

- Task/Procedure/Material

- Context
  - (situation, meanings, functions)
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- → speaker-specific coding and variation must be taken into account

- For example, it is known across languages (e.g., English [intonation] Mandarin [tone] and Swedish [tone accent]) that **female speakers** can bridge **F0 distances faster** than male speakers (Sundberg 1979; Xu and Sun 2002; inter alia)

- Female speakers have a different **glottalization behaviour** and can hence show a different interplay of glottalization with other related cues in the production of low tones, phrase boundaries, disfluencies etc.; The “modal” voice of female speakers is typically breathier than that of male speakers (Klatt & Klatt 1990; Simpson 2010; inter alia)
Part 2 « Theory-oriented Speech-Data Acquisition »

• Informants/speakers are not merely generators of a speech signal!
• Speaker-specific variation and strategies must be taken into account.

• Peters (1999, 2000) and Ambrazaitis (2005) found gender differences in the realization of terminal F0 falls at the ends of utterances in German and – more recently – also in English and Swedish. Compared with male speakers, female speakers prefer “pseudo terminal” falls that end in a deceleration and a slight, short rise (2-4 st) at a relatively low intensity level.

• These “pseudo terminal” falls may in extreme cases be confused with actual falling-rising utterance-final intonation patterns (rise typically > 6 st). However, they differ from rising-falling patterns in terms of both phonetic form and communicative function.
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- → speaker-specific coding and variation must be taken into account

Communicative function that may be assumed cross-linguistically (in line with the frequency code):

Compared with a terminal fall, a pseudo-terminal fall reduces the dominance of the speaker and/or the finality of the statement

Kiel Corpus: „Dienstag wieder frisch gebrannte Mandeln“

26.05.2012

Oliver Niebuhr & Alexis Michaud
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- \(\rightarrow\) speaker-specific coding and variation must be taken into account

- But gender differences are of course not the whole story.
- Niebuhr, D’Imperio, Gili Fivela, Cangemi (2011) found in a large cross-linguistic study on German and varieties of Italian that speakers use different strategies to realize the contrast between low and high pitch accents like: \(H+L^{*}\) vs. \(H^{*}\) or \(L+H^{*}\) vs. \(L^{*}+H\)

- Some speakers change the temporal coordination between the F0 peak and the associated syllable = „Aligners“
- Other speakers change the internal timing of the F0 peak and keep its alignment constant = „Shapers“
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- → speaker-specific coding and variation must be taken into account

- However, most speakers made use of both peak alignment and peak shape to different degrees.

- **German:**

![Graph showing 'Shapers' and 'Aligners']

26.05.2012

Oliver Niebuhr & Alexis Michaud
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- → speaker-specific coding and variation must be taken into account

- However, most speakers made use of both peak alignment and peak shape to different degrees.
- **Neapolitan Italian:**

![Diagram showing latencies for Neapolitan Italian with labels "Aligners" and "Shapers".
Part 2 « Theory-oriented Speech-Data Acquisition »

- Informants/speakers are not merely generators of a speech signal!
- → speaker-specific coding and variation must be taken into account
Informants/speakers are not merely generators of a speech signal!

Speaker-specific coding and variation must be taken into account.

A number of studies showed additionally that factors like age, smoking and drinking habits, weight, social class, musical training etc. also affect prosodic patterning in general and tone/intonation patterning in particular.

In consequence:

- select your speakers/informants carefully and try to balance known speaker-related factors (i.e. in order qualify for a specific recording/corpus “normal speaking/hearing abilities” may not be the only criterion);
- make your speaker sample as large as possible (e.g., 4 speakers do not represent “the language xy”)
- let your speakers/informants fill out questionnaires that collect as many meta data as possible (e.g. more than the usual suspects: age, gender, home town)
- During analysis, compare within subject means and – if necessary – create sub-samples before you calculate overall means for each measure
Part 2 « Theory-oriented Speech-Data Acquisition »

Question / Aim
- documentation
- investigation

Subject
(grammer, phonology, sounds, prosodies)

Informants/speakers

Speech data acquisition

Context
(situation, meanings, functions)

Task/Procedure/Material
Part 2 « Theory-oriented Speech-Data Acquisition »

• Avoid monotonous tasks!

• Especially in experiments with a larger number of cross-combined independent variables it is often necessary to elicit numerous target (e.g., words, pitch accents, etc.) with specific phonetic properties in prosodically controlled environments.

• Frequent method: speakers read lists of similar sentences
  • Constant carrier sentences like “I don’t know the word ___”; “The next word is ___”; “I have seen ___ on the table”; …
  • Carrier sentences with variable wordings and constant syntactic structures like “The house is on the mountain”, “The plate is on the table”, “The dog lies on the floor”, … (= NP_{[ART,N]} + VP_{[Pres.Sg.]} + PP_{[Prep,ART,N]})

• Such carrier sentences allow for a maximum degree of control, but represent a strong abstraction from everyday communication, which limits the generalization of the analyzed data.
Part 2 « Theory-oriented Speech-Data Acquisition »

- Avoid monotonous tasks!

- If you cannot avoid using the sentence-list method, be aware that
  
  …presenting the entire list at once to the readers (on a single sheet of paper) may create artefacts in the form of “list intonations”

  ![Diagram](image)

  - → possible solution: add dummy sentences at the beginning and end of the list
Part 2 « Theory-oriented Speech-Data Acquisition »

• Avoid monotonous tasks!

• If you cannot avoid using the sentence-list method, be aware that

• …readers are not “mere speech generators”. They may spontaneously establish **semantic/pragmatic relations** between individual sentences, even if these sentences are on separate sheets of paper
  
  • For example, “Peter came by car” → “Meghan came by bus” → “Steve came by boat” may cause that “bus” and “boat” are realized with prosodies of contrastive topic = prosodic artefacts.

  • Similarly: “The next word is house” → “The next word is window”;

  • “The plate is on the table” → “The glass is on the table” (“table” becomes given information; “glass” is realized in contrast to “plate” = prosodic artefacts)

• → possible solution: randomize the sentences differently across the speakers. Between sentences: use syntactic constituents – particularly target words – that are as unrelated as possible (vary functions words like prepositions, pronouns, etc. if possible)
Part 2 « Theory-oriented Speech-Data Acquisition »

• Avoid monotonous tasks!

• If you cannot avoid using the sentence-list method, be aware that
  • …the “instrument of spoken language” can become “blunt” if your list contains too many sentences
  • That is, speech production becomes a mere muscular exercise that is decoupled from its original aim of conveying a message or getting the dialogue partner to do something
  • Consequence: prosodic patterns become monotonous and may also be artificially stabilized/stylized by speakers = prosodic artefacts
  • For example: Preliminary data of an ongoing study at the University of Kiel
    • Speakers (12 so far) read individually randomized lists of 200 sentences
    • 2 lists: one with sentence of the type “The next word is ___” and another one with sentences of the type NP-VP-PP (“The cat sleeps on the sofa” etc.)
Part 2 « Theory-oriented Speech-Data Acquisition »

- Avoid monotonous tasks!

- Results:

  Pitch level and pitch range decrease across the 200 read sentences

Speaking rate increases, variation in pitch-accent alignment decreases

Why? Training or functional erosion?
Part 2 « Theory-oriented Speech-Data Acquisition »

• Avoid monotonous tasks!

• As regards the (preliminary) results for pitch-accent alignment, it should be noted that most studies on the well-known phenomenon of “segmental anchoring” are based on lists of read sentences → To what extent is “segmental anchoring” facilitated by the elicitation method?

→ possible solution: as the largest prosodic changes seem to occur between sentence #50 and #100, try to use lists of about or less than 50 sentences (at least per session). Avoid multiple repetitions of sentences in a reading session.
Part 2 « Theory-oriented Speech-Data Acquisition »
The segmental “layer” cannot be analyzed separately from the prosodic “layer”!

Sound segments $\rightarrow$ F0 contours

We know from tone-accent and pitch-accent patterns that the segmental make-up of the corresponding syllables affects the alignment of F0 contours; these **push or pull effects** are similar across languages (Ladd 2003, 2008)

In simple terms:
- F0 valleys seem to be more consistently aligned than F0 peaks
- F0 peaks move to the **left** in closed syllables, particularly in those closed syllables with obstruent(s) in the syllable coda
- F0 peaks move to the **right** in open syllables, in syllables with long vowels and in syllables with obstruent(s) in the syllable onset
Part 2 « Theory-oriented Speech-Data Acquisition »

- The segmental “layer” cannot be analyzed separately from the prosodic “layer”!
- F0 contours ← Sound segments

- By variation in the spectral-energy distribution, each sound segment has the potential to create or to contribute to a particular pitch impression.
- At least in German, the spectral-energy levels and distributions of sound segments vary in such a way that they match with the F0 context → “segmental intonation” (Niebuhr 2008, 2009, 2011)… though to a speaker-specific degree!

- For example, voiceless obstruents sound higher at the end of final F0 rises and lower at the end of final F0 falls.
Part 2 « Theory-oriented Speech-Data Acquisition »

- The segmental "layer" cannot be analyzed separately from the prosodic "layer"!
- F0 contours ← Sound segments

- Similarly: in context of F0 rises the spectral transition in closing diphthongs (/ai/, /au/) starts earlier and ends in a higher vowel quality
Part 2 « Theory-oriented Speech-Data Acquisition »

- The segmental “layer” cannot be analyzed separately from the prosodic “layer”!
- F0 contours ↔ Sound segments

- At the ends of utterances, the realization of low-falling F0 movements can be in conflict with the production of final voiceless obstruents
  - → The degree to which the F0 fall
    - is truncated by the final obstruent
    - or compressed into the preceding voiced sound segments

- seems to vary with the ability of the final obstruents to convey “segmental intonation”

- → For example, sibilants [s,ʃ] stimulate truncation, whereas [f] and [h] stimulate compression of the final F0 fall (cf. Pfitzinger & Ohl 2009).
The segmental “layer” cannot be analyzed separately from the prosodic “layer”!

F0 contours ↔ Sound segments

The realization of pitch accents can be linked with specific duration and intensity patterns in and around the accented syllable

→ Pitch accents can create characteristic imprints in the segmental “layer”
→ Individual “micro rhythms” that seem to support the perception of the phonological features of the pitch accent.
The segmental “layer” cannot be analyzed separately from the prosodic “layer”!

In consequence:

- Consider that the measures of your prosodic/intonational analysis are additionally shaped by the underlying string of sound segments.
- Consider that the measures of your segmental analysis are additionally shaped by the superimposed prosodic/intonational patterns.
- Keep in mind that:
  - Voiced sound segments, including vowels and diphthongs, may also be involved in “segmental intonations”.
  - Using completely sonorant utterances are not a way to separate the 2 “layers” → similar to obstruents, sonorant consonants also cause mircoprosodic F0 perturbations; sonorants, particularly approximants, can obscure segment boundaries you may need as references in prosodic analysis.
Part 2 « Theory-oriented Speech-Data Acquisition »

- Question / Aim
  - documentation
  - investigation

- Subject
  - (grammar, phonology sounds, prosodies)

- Speech data acquisition
  - Informants/speakers
  - Task/Procedure/Material
  - Context (situation, meanings, functions)
Part 2 « Theory-oriented Speech-Data Acquisition »

- When you elicit data, you initiate a communication process; and communication is a major way of social interaction → Be aware that you always elicit more than the forms and functions/meanings that you aim at!

- Sociophonetic and sociolinguistic studies agree that dialogues between 2 men, 2 women or a man and a woman differ in many respects, such as wording, F0 register and range, phrasing, turn-taking/yielding, speech reduction etc. (cf. Giles 1991 inter alia)

- Similarly (or maybe for the same reason!), the linguistic and phonetic patterns of speakers in dialogues are shaped by social or cultural hierarchies as well as by familiarity with the dialogue partner (cf. Campbell & Mokhtari 2003; Coates 2004, inter alia)
Part 2 « Theory-oriented Speech-Data Acquisition »

• When you elicit data, you initiate a communication process; and communication is a major way of social interaction → Be aware that you always elicit more than the forms and functions/meanings that you aim at!

• It even matters at which time of day the recording takes place. Görs (2011) created a large corpus of more than 30 German speakers, who read texts (1) early in the morning, (2) at about noon, and (3) late in the evening. She found systematic prosodic differences as a function of the time of day:

• In the morning, speakers show a slower speaking rate and a lower F0 level as well as stronger glottalization at prosodic boundaries

• Speaking rate and F0 level increase at noon; the same applies to the level of speech reduction

• In the evening, the F0 level is lower again; the speaking rate remains high, but with fewer speech reductions; the voice quality is overall breathier; the speech rhythm is most pronounced (in terms of prominence cues)
Part 2 « Theory-oriented Speech-Data Acquisition »

• When you elicit data, you initiate a communication process; and communication is a major way of social interaction → Be aware that you always elicit more than the forms and functions/meanings that you aim at!

• Finally, it must be taken into account that different **elicitation strategies** represent different communicative frameworks

• → Elicitation tasks are not simply exchangeable, and the data they yield are not necessarily comparable across studies

• For example, the difference between broad-focus accents and narrow/contrastive-focus accents has been elicited - among others - by
  • read monologues (e.g., text passages)
  • read short A-B dialogues in which B responds to either a question or a statement of A
  • unscripted dialogues, recorded in a cooperation-task scenario (e.g., ‘Maptask’)
Part 2 « Theory-oriented Speech-Data Acquisition »

• When you elicit data, you initiate a communication process; and communication is a major way of social interaction → Be aware that you always elicit more than the forms and functions/meanings that you aim at!

• Görs and Niebuhr (2012) compared the three elicitation tasks (read monologues, read dialogues, spontaneous dialogues) with the same 8 speakers and the same set of target words that were produced with broad and narrow/contrastive focus.

• Results:
  • In read monologues, the difference between broad and contrastive focus is only a matter of intonation → compared with broad-focus, narrow/contrastive-focus target words show longer and slightly higher F0 movements.
  • When a real dialogue partner is present, the intonational difference changes from a predominant alignment to a predominant scaling difference. Moreover, narrow/contrastive target words show additionally a higher, steeper intensity increase and a lengthened syllable (onset), even more so when communication serves to solve a joint task (as in ‘Maptask’).
Part 2 « Theory-oriented Speech-Data Acquisition »

• When you elicit data, you initiate a communication process; and communication is a major way of social interaction → Be aware that you always elicit more than the forms and functions/meanings that you aim at!

• Interpretation:

• When a dialogue partner can be addressed (= not in monologues), the actual intonational signalling of broad vs. narrow/contrastive focus is **overlaid** by a type of **emphatic accentuation** (‘reinforcement’), whose function is to highlight the truth value of the target words

• Reinforcing information is even more important when a joint task must be solved (= difference read vs. spontaneous dialogues)

• In consequence: Choose your elicitation task carefully → What kinds of functions/meanings do you want to elicit? What kinds of functions/meanings do you want to exclude/control? → Which communicative task/framework meets your requirements (at least approximately)?
Part 2 Summary

- Informants/speakers are not merely generators of a speech signal
- Avoid monotonous tasks
- Sounds segments cannot be analyzed separately from prosodies
- You elicit more than the forms and functions/meanings that you aim at
- The importance of a well-conceived speech-data acquisition is often underestimated…
  - not so much because researchers select the elicitation method according to personal preferences and experiences
  - but mainly because the formal and functional complexity of human speech is underestimated

Speech

- Conveys information, initiates actions of the dialogue partner
- Facilitates social interaction, creates identity

Can we exploit our corresponding knowledge as a means of exper. control?
Part 2 Summary

- An unspecific collection of speech data, irrespectively of its size, is usually not very helpful in answering specific research questions.
  - “it is not economical and practical to make a [...] corpus all-inclusive and all-embracing” (J. Xu)
  - “even in a very large corpus of [...] hundreds or even thousands of utterances [...] it will be very difficult to find a set of four to ten versions of the ‘same’ utterance” (Himmelmann 2006:169)
  - “We cannot hope to anticipate all future needs” (Mithun 2001:53)
- However, an unspecific data collection, particularly spontaneous speech, can be a great source of observations and ideas, and a useful reference to support results from purposefully recorded corpora
- Investigating specific research questions requires a separate, target-oriented elicitation of speech data
  - In order to be comparable “utterances have to convey the same meaning and, most importantly, they have to be performed with the intention of achieving the same illocutionary act” (Himmelmann 2006:168)
Part 2  Summary

- If the function/meaning that you aim at is not interactional, you may want to elicit monologues (note: sentence mode is interactional)
- In all other cases, you should use dialogues

Unscripted starting-point

Written texts as starting-point

Pre-select a sufficient number of speakers and collect a wide range of personal data

facilitate a speaking style by varying the age, gender, familiarity, social status, recording experience etc. of your dialogue partners

Producing spontaneous speech in the lab as well as producing spontaneously-sounding read speech **both** require a certain extroversion, fluency, language competence, self-confidence; pre-select your speakers accordingly
Part 2  Summary

• If the function/meaning that you aim at is not interactional, you may want to elicit monologues (note: sentence mode is interactional)

• In all other cases, you should use dialogues

Ways of including key words in unscripted speech

- Maptask → A explains a path on a map to B, who must trace the path (sometimes in a given amount of time): icons on the map and the directional instructions function as key words (Anderson et al. 1991)

- Videotask → A and B view two slightly different excerpts of a well-known TV show/series. Then, they have to discuss in order to find out the differences: major elements of the show/series (names, places, items etc.) function as key words (Peters 2001, Kohler 2007)

- Appointment-making task: A and B have to make appointments for business or leisure events. Days, times, places etc. are key words (Kohler et al. 1997)

- Retelling of a story or of a comic strip (Mosel 2006)
Part 2 Summary

- If the function/meaning that you aim at is not interactional, you may want to elicit monologues (note: sentence mode is interactional)
- In all other cases, you should use dialogues

Way of eliciting segmentally and prosodically controlled, but spontaneously-sounding speech from written texts:
- Create dialogue texts on everyday topics; if possible, integrate common reduction phenomena in the orthographic representation
- Let your carefully selected and paired dialogue partners practice the texts in advance; allow them to adjust the texts slightly to their own way of expression by introducing, omitting or replacing words and phrases
- **Elicit specific prosodic/intonational patterns by creating adequate semantic/pragmatic contexts in the preceding utterance(s)**
Part 2 Summary

- If the function/meaning that you aim at is not interactional, you may want to elicit monologues (note: sentence mode is interactional)
- In all other cases, you should use dialogues

Way of eliciting segmentally and prosodically controlled, but spontaneously-sounding speech from written texts:
- Instruct the speakers to judge each others production performances and to repeat the dialogue until they are both satisfied and agree that they have produced an everyday-sounding dialogue
- Advantage: allows for a high degree of segmental and prosodic control. If the speakers are well selected, the output is in many prosodic respects very similar to that of actual spontaneous speech.
Thank you for your attention
Encountering an opposition between /dzi/ and /dz/, /tɕi/ and /tɕw/, /tɕʰi/ and /tɕʰw/. Both sets are apicalized to some extent. (Language: Yongning Na)
• Sufficient for some research and teaching purposes
• Useful in fieldwork: further verification of data (exploratory value)