Markedness Tutorial

Paul de Lacy
3/16/2019

This document accompanies a PowerPoint presentation. The numbers below refer to the slides.

#1. Title slide

My name is Paul de Lacy. My talk is the first of two tutorials about markedness.

I believe I am supposed to present a talk, and then have 15 minutes for questions. Are people allowed to ask me questions during the talk?

- I have uploaded the presentation and my talk AND references I cite to this URL. You are welcome to go there now or later.

TIME: 30s

#2. Goals

The goal of this tutorial was identified by Tobias Scheer.

- It is about **markedness and melodic primes, or features.**
- Are they universal, or not? And innate or not?

- Tobias asked me to argue for the 'classical' position: markedness and features are **universal** and **innate**.
  So, the goal of this tutorial is to address two questions:
- Is markedness universal?
- Is markedness innate?
  - Markedness could be universal and *not* directly innate. This would be a situation where every human uses functional algorithms to determine markedness relations, and the results are always the same due to the nature of the algorithms.
  - I am not going to talk about features directly. I will talk about features as they relate to markedness. I'm happy to answer questions about features, though.

TIME: 1:30 (+1)

#3. Goals ctd

These questions are difficult to answer. So, I have broken down the problem into several pieces.

- What is markedness?
- Where does it apply?
- What is a theory of markedness?
- Does markedness *require* universality?
- How do we discover markedness? Methodology.
- What evidence is there for markedness?

- What are the current issues in markedness?

- I am going to argue that:

- We need a clear theory of phonological markedness before we can ask any questions about it.

- We need to examine the theory carefully to see what it says it is and is not responsible for.

- We need to evaluate our methods for gathering evidence.
    - So far, we have not.
- We need to evaluate our evidence carefully
    - So far, we have not.
- Do we know anything about markedness?
    - No, but we *could* if we tried.

TIME: 3:45 (+2:15)

#4. A. What is markedness? Non-cognitive tradition

(1) *"Markedness" is ambiguous*
(a) Markedness in language taxonomy
(b) Markedness in cognition

(2) *Taxonomic markedness*
(a) The goal: To describe and classify languages (and potentially parts of languages)
(b) For Taxonomists, Markedness is a way of classifying and describing.
(c) American Structuralism, Greenbergs universals, *the Prague School*.
(d) For example, Zellig Harris, in talking about his system of language description said "{READ ALOUD}"
(c) Concept #1: *Language* – needs a definition!
(d) Concept #2: *Markedness* – A way of classifying languages – languages with highly marked inventories vs. less marked inventories, etc.
(e) Definition? There are hierarchies (of segments/features?), and these hierarchies are reflected in languages in quantitative ways: by having more unmarked segments in inventories, by unmarked segments being the outcome of neutralization, and potentially other things.

TIME: 4:45 (+1:25)

#5. A. What is markedness? Non-cognitive tradition 2

- As an example, we can observe that we find [t] in many languages – over 99% of them in UPSID.
- In contrast, we don't find many [q]s – voiceless uvular stops (13.8% in UPSID).
- Therefore, [q] is more marked than [t]
- We can do this for all consonants and vowels, and then build up a concept of 'more or less marked inventory'
- We might even observe that we only need two categories of language in terms of [t] and [q]: i.e. languages with **both** [t q], and languages with [t] and not [q]. We won't find languages with [q] and not [t].

- In other words, if you have something marked in your inventory, you'll have **everything less marked, too.**
- We can see that markedness is a useful language classification tool, then.

Time: 5:45 (+1:00)

#6. A. What is markedness?

There is another meaning of 'markedness'.
It is a *very different* meaning from the taxonomic meaning, though the meanings are related.

- But firstly, we must ask 'which cognitive tradition'?
    - I am focusing on Generative theories. These start with Chomsky & Halle (1968) – SPE.
    - There have been many markedness theories since then. My own is in de Lacy (2002) and (2006).
    - I am going to try to be as general as possible. However, at certain points we will **need** to talk about a specific theory, so I will talk about de Lacy (2006)'s.
- OK, In general, what does 'markedness' mean for Generative theories of the phonological module?
- To answer this question, we need to define **state**

TIME: 6:45 (+1:00)

#7. A. What is markedness?

- A PhM 'state' is a complete input-output mapping.
- It's the set of all possible inputs and their output mappings.
- Here's an example. Input 1 maps to Output 1, Input2 deletes completely – it maps to null. Inputs 3 and 4 map to the same output, and so on, into potentially infinite inputs.
- Of course, there are many possible PhM states because we see many different PhMs.
- This term 'state' simply describes the basic idea behind a Generative Grammar. States are part of every Generative theory, from SPE to Optimality Theory, and all others.
- You might be more familiar with the term 'grammar'. A 'grammar' is a description of a state, using rules, constraints, filters, or whatever.
- I will use the term 'state' here because I want my comments to apply to as many Generative theories as possible.

TIME: 7:30 (+0:45)

#8. A. What is markedness?

The big question about states:
Are there asymmetries in states?

- Are there some states that we can define, but DO NOT exist.
    - In other words, the PhM cannot generate them.

- For example, suppose we find that there are states in which /p/->[t] in codas, but no state where /t/->[p]. This is a state 'asymmetry'.
    - Of course, our PhM theory will need to account for such state asymmetries.

- Now, here's the key point: A theory of state asymmetries **is** a theory of markedness.

- From a formal point of view, ‘markedness’ is simply a way to refer to state asymmetries.

TIME: 8:15 (+0:45)

#9. A. Taxonomic vs. cognitive

Now that we have defined some formal terms, we can ask "What's the difference between taxonomic and cognitive markedness?"

- There are huge differences, but let's focus on two.

- Taxonomic markedness is about "languages"
  - A 'language' is a social concept – a system of communication used and shared between members of a community.
- Cognitive markedness is about "states"
  - In other words, which states do and do not exist?
- A major point of confusion here is that phonologists often use the term 'language' instead of 'state', or 'grammar'. I will no doubt do this today. Please forgive me.

- The second difference is about generality.

  - Taxonomic theories see a 'language' as general and amorphous. Consequently, everything about a language can fall under the same classification system – i.e. 'taxonomic markedness'.
  - Cognitive theories are very specific about what does what. In this case, we're only interested in the PhM.
    - We must explicitly, and rigorously exclude other modules and mechanisms from our purview.

TIME: 9:15 (+1:00)

#10. A. OT 'markedness'

By the way, there's another use of 'markedness'. Optimality Theory's output constraints are called 'markedness constraints'. This is just bad terminology. Markedness is about asymmetries – i.e. both faithfulness and 'markedness' constraints.

However, the term is here to stay. What a shame.

TIME: 9:45 (+0:30)

#11. A.Summary

Read the slide.

TIME: 10:45 (+1:00)

#12. B. where?

In taxonomic markedness, the focus was largely on 'surface' properties.

Why? Because major descriptivist systems had only one level – a surface level.

Phoneme theory is a very big exception to this, of course.

However, the focus on surface description can still be felt in cognitive theories. So much so, that we often forget to look at all of the potential places for asymmetries in a grammar.

So where else should we look?

TIME: 11:15 (+0:30)

#13. B. where?

There are 3 major places.

- The input
- The output
- The mapping between the input and output
- We could in principle find markedness effects in each of these areas. Let's see how.

TIME: 12:00 (+0:45)

#14. B. where? Output

Let's start with the easy case: the output.

An output markedness statement would be "No output has output o_1 in the context of {other state properties}".

- For example, "No state has trisyllabic feet". Or, more clumsily, "For all states, there is no output with a foot that contains three syllables."
- "If a state has no [t] in any output, then it has a glottal stop in some of its outputs."
    - Or, "If a language doesn't have a [t], then it has a glottal stop."
- Lots of the focus in markedness has been on outputs and output statements like these. But is this really where the action is?

TIME: 12:45 (+0:45)

#15. B. where? Output

Here's a table that shows the stops in syllable codas in a variety of 'languages', or 'states' for us.

- The interesting thing here is that you can get almost anything you can imagine.
- We'll talk more about place of articulation markedness later. However, you can immediately see that there is no evidence here for any asymmetries. There aren't any: everything is attested. Almost.
    - We get systems which have everything except for [k]
    - And ones with everything except for [p]
    - And ones with everything except for [t]
    - And ones with everything except for glottal stop
    - And every other gap you can imagine.
    - Except for two, which turn out to be significant later on.
- The point is that even though it is *tempting* to focus on the output in our search for markedness, it turns out that it's not a good place to look. In fact, it's almost useless, and we'll see why in a moment.

TIME: 13:45 (+1:00)

#16. B. where? Input

What about the input?
An input asymmetry would be one in which there are restrictions on input strings relative to, say, output strings: in other words, can inputs contain everything that outputs do?

Interestingly, the answer is 'no'. Inputs do not contain prosodic structure, otherwise we would see contrastive syllabification and footing.

We **do** see contrastive main stress, but see some of my recent work for discussion about that.

So, our PhM theory needs to account for input asymmetries.

However, these kind of asymmetries are not usually what people think of when they talk about markedness. So, let's turn to mapping asymmetries.

TIME: 14:30 (+0:45)

#17. B. where? Mappings

- A mapping asymmetry involves situations in input-output mapping where /A/->B, but B never -> A, usually in a specific environment.
- A classic example involves coda voice neutralization. We see [+voice] map to [-voice], but not [-voice] -> [+voice]
- A more complex example involves place of articulation in stops.
- Of course, /k p t ?/ all map to themselves. That's boring, so let's get rid of that.
- More interestingly, in non-assimilatory and non-dissimilatory environments, coda /k p/ can map to [t].
- They can also map to [?], and so can /t/.
- And that's it. There are no mappings of /t/ to [p] or [k], or /k/ to [p], or /p/ to [k], and so on. These are **mapping asymmetries**, then.
- By the way, what's *really* interesting about these asymmetries is that you can generate almost any *output coda inventory* with them.
    - For example, do you want a [p t ?] inventory? No problem: {Click}
    - Do you want an inventory that lacks the low markedness [t]? No problem: {Click}
    - Amazingly, even though we have many PoA mapping asymmetries, we have almost NO PoA output asymmetries. A striking result.

TIME: 16:00 (+1:30)

#18. B. where? Mappings

Here's another example: voicing in stops in codas.
Focusing on states that lack assimilation or dissimilation in the appropriate environments…

We find only two states here: either voiced stops in codas, or voiceless stops.

We don't find states in which some voiced stops devoice and others remain voiced.
Here is an interesting situation where the extreme paucity of possible mappings results in easily detectible output asymmetries.

So, markedness in mapping is where we find the most asymmetries.
Asymmetric mappings do not necessarily imply asymmetric outputs, so it's not a great idea to focus on outputs.

The catch: It's *difficult* to study IO mappings. You need a special kind of evidence (morpho-phonological alternations) that's hard to find. It's easy to flip open a grammar and find the consonant chart. It's harder to open a grammar and find all the possible stop neutralizations. Often the grammar doesn't tell you about alternations, or tells you the wrong thing. Often you need a dictionary that shows allomorphs, too. It's difficult and time consuming. However, it's also necessary.

TIME: 17:00 (+1:00)

#19. B. where? Representation and computation

Another issue is where markedness should reside in one's theory: in the representation or computation.

'Markedness in computation' means that there are computational processes - mappings, as imposed by rules, constraints, etc. – which involve /A/->B but not /B/->A.
Like, say, dorsals becoming coronals in codas, but not coronals becoming dorsals.

'Markedness in representation' refers to a property of certain representations that allow them to undergo some process, while other representations do not. The classic case involves complexity, where general cognitive processes will simplify complex representations, but leave simple ones alone (because they're already simple). In this case, markedness is implemented through reduction of complex structure.

TIME: 18:00 (+1:00)

#20. C. Principles of Markedness

Now that we have some idea about what a theory of markedness should do, and where we should see its effects, we can fruitfully talk about an actual theory of markedness – de Lacy (2006)'s theory.

In this theory, markedness hierarchies are orders between feature values.

- For example: +voice **is more marked than** -voice
- Dorsal > labial > coronal > glottal, where each is seen as a single feature [Place] with multiple values.

There are more complex hierarchies, too, involving environments in a predictable way.

- So, we get an onset sonority hierearchy.
- And a foot head – stressed syllable – sonority hierarchy.

- The usefulness of these hierarchies is that they determine the form of constraints – both output and faithfulness, in OT terms. In this way, they can affect the grammar: they define possible states, and that's why these hierarchies are the core of a theory of markedness.

- Crucially, markedness hierarchies are **feature** hierarchies, not **segmental** ones.
    - There is therefore no meaning to the term "Glottal stop is the least marked consonant", or "[d] is less marked than [k]". It all depends on which hierarchy we're talking about.
    - A big consequence is that two segments could be both less or more marked than each other depending on the hierarchy. [b] is more marked than [k] on the voicing hierarchy, but less marked on the PoA hierarchy. Glottal stop is least marked in PoA, but highly marked in onset sonority.

TIME: 19:15 (+1:15)

#21. C. Principles of Markedness: Preservation of the marked

- While a lot of the focus in markedness used to be on the surface, a theory of state asymmetries has to look at mappings.
- That's what 'Preservation of the Marked' is about: A theoretical mechanism that preserves very marked elements in a hierarchy.
- Dorsals are the most marked PoA feature. So, they are both subject to specific restrictions – *dorsal, and special preservation – FAITH[dorsal]
- We see here a specific implementation in Optimality Theory.
- Faithfulness to dorsals is most important.
- Otherwise, there is pressure to eliminate marked elements, and this pressure outranks all preservation.
- The result is a system where the most marked element in terms of PoA – the [k] – survives, but less marked elements neutralize to the least marked element.
- Again, we end up with a 'marked' inventory, but entirely through constraints that are asymmetric in their form.

TIME: 20:45 (+1:30)

#22. C. Principles of Markedness: Conflation

The final big principle is that of conflation.
Is it right to say that A is more marked than B, or that B is never more marked than A?
In the latter case, we allow for situations where A and B are treated as having the same markedness in a language.
The arguments are complex, but I argue that it is the latter.
Perhaps we should use the 'greater than or equal to' sign rather than the 'greater than' one.

TIME: 21:45 (+1:00)

#23. C. Principles of Markedness: Summary

I have just summarized – very briefly – the theory of markedness presented in de Lacy (2006)'s theory.
Markedness hierarchies refer to features, not segments.

So, we can have apparent contradictions where a segment is less marked on one scale, and more marked on another, relative to some other segment.

We have markedness reduction, where marked elements are mapped to less marked ones.

We have markedness preservation, where more marked elements survive otherwise general processes.

We have markedness conflation, where markedness categories can be treated the same for a specific language.

TIME: 22:30 (+0:30)

#24. D. Universal and innate?

- Why did I spend all this time talking about the formal properties of a theory of markedness?
- Well, if we didn't establish *what markedness means*, we cannot ask the question we want to ask:

So, what does the question "Is markedness universal and innate?" mean?

- In the de Lacy (2006) theory, it means:
  Are feature value hierarchies and the mechanisms that refer to them (constraints, in OT) universal and innate?

- The answer must be yes. If it was not, there would be no universality of consequences.
- I suppose they might be universal and *not* innate *if* we create a theory that has a mechanism for learning hierarchies and features, and it always converges on the same hierarchies and features. However, outside of this very specific – and surprising – situation, universality of consequences means universality of formalism.

- We can also now ask other useful questions:
  - Can we have markedness without universal features?
  - Of course, 'markedness' here refers to PhM state asymmetries.
  - The answer, set in de Lacy (2006)'s theory is *no*.
  - Why? Because the question is, restated in present terms, "Can we have universal feature value hierarchies without universal features?"
  - No! Because hierarchies cannot refer to features without first having those features.
- Finally, if we have universal features, *must* we have universal markedness?
- The answer is *no*. It's possible to have a theory that just has features and no feature value hierarchies.
- However, the bigger question is "Can we formally define a theory that has no state asymmetries?"
  - That's actually hard once you go beyond trivial theories. It's easy for state asymmetries to occur if you limit the input, input->output mapping, or output in any way.
  - So, if anyone says to you "I believe that the phonological module can do anything.", you should ask them what they mean. What is their formal theory of the phonological module?

TIME: 24:00 (+1:30)

#25. E. Methodology

I have just provided a technical answer to the question "What does it mean for markedness to be innate?" After defining what markedness is, we were able to ask under which conditions could it be innate.

In a sense, we have answered the questions "Is it possible to formally define a theory that produces PhM state asymmetries?" The answer is yes.

We have also answered the question "Is it possible to have markedness but no universal features?" The answer, in relation to de Lacy (2006)'s theory, is no.

The obvious next question is "What evidence is there for this theory? What is the evidence for a theory of markedness like de Lacy (2006)'s?"

To answer this, we need to know "Where do we get our evidence for the PhM from?"

The answer is that we get them from 'devices'. Devices such as humans, as in fieldwork, from spectrograms – via humans or software algorithms, or from grammars, or rather from human interpretation of grammars, and, and so on.

But there is a problem with all of these devices. None of them have *direct access* to the PhM. To either it's inputs, mapping, or outputs.

TIME: 25:00 (+1:00)

#26. E. Grammar structure

Let's take a closer look.

Here is a Generative model of the speech mechanism.
The PhM has several inputs: the lexicon, syntax, and morphology. These may be ordered – it doesn't matter for today.

The PhM outputs to a phonetic module, along the lines of Keating, Kingston, Cohn, and so on.

The phonetic module outputs to an articulatory interface, which then motivates articulators. These can produce sound, under the right conditions.

We really have two kinds of non-human devices for studying the PhM. One measures articulators. For example, palatography. The other – much more common – measures perturbations in the transmission medium – i.e. speech sounds. This is *Praat*, spectrographs, pitch trackers, and so on.

Notice, however, that all of these devices are *very far* from the PhM. The Phonetic module and neuro-motor interface intervenes. Also, there are modules that *precede* the PhM, and the PhM is inevitably influenced by them.

TIME: 26:00 (+1:00)

#27. E. Responsibility

- So, whenever we look at the output of our devices, we *cannot assume* that they reflect the PhM *alone*. In fact, we know that they **do not**. We must *always* ask ourselves: which elements of my observation is the PhM responsible for? Which is it not, or potentially not?
- For example, suppose we detect speech sounds [pa̰], with creak on the [a], and no full glottal closure.
- We must ask ourselves: Which part of the speech production process is responsible?
- It is possible that the PhM has a coalescence process – option A.
- But it is also possible that it was the phonetic module that implemented – perhaps optionally – the glottal stop purely as creak on the preceding vowel.
- Finally, it is possible that the phonetic module told the articulators to have a full glottal closure. However, the articulators misfired and failed to implement the complete glottal closure command.

- Any of these are options in a sophisticated model of the speech production mechanism.

TIME: 27:00 (+1:00)

#28. E. Responsibility

There are also considerations that are **far outside** humans. This is best explained through an example.

It's common to see statements like this "[s] is found in 80% of the world's languages, and so [s] is quite unmarked."

- What are we to make of such a statement?
- Well, firstly, it is said within a taxonomic framework, not a cognitive one, so we will need to translate it.
- We need to express it in terms of the PhM and states.
- And what we can extract are two statements.
- One is that there is some PhM state that has an output with [s] in it.
- The other is that there is some state that has no outputs with [s] in it.
- We arrive at this conclusion by making a logical leap. That leap is that if there is a "language" with [s] in it, then there is a human somewhere who has an L1 intact PhM state that produces outputs with [s]. This seems like a reasonable assumption. It's not, by the way, but let's assume it.

TIME: 27:45 (+0:45)

#29. E. Responsibility

- In general terms, we can say that if we observe a human to map /A/ to [B] in their PhM, then we can conclude that there is a PhM state in which /A/->[B].
    - This is obvious, of course, but is worth stating because it is sometimes ignored.

- For example, way back in the past, back when people believed that coronals were unmarked, I would often hear something like this "Every language has a [t], except for Hawai'ian." And then after that we would assume that Hawai'ian speakers didn't exist.
- Of course, Hawai'ian speakers, assuming PhM homogeneity, mean that we need states with [t], and states without [t], and that's all we can conclude.

- Can we now go further? Could we infer something like this:

    - "80% of PhM states have some output that contains [s], but 20% of PhM states have no [s]-containing output."
    - Actually, no! We can't.
    - For this statement to be true, we would have to assume that currently existing languages perfectly reflect the capabilities of the PhM – i.e. that they are somehow a perfect sample of PhM states.
    - However, this is completely implausible. For a start, we would have to say that the languages we have *now* are a perfect sample, and not in the past or in the future.
    - We would also have to ignore all the capricious influences on language: war, disease, and other social processes like immigration. These have had profound effects on languages, and so the idea that we have a perfect sample of PhM states is really unbelievable.

- I really believe that percentage statements involving inventories are not probative in regard to the PhM. And that includes statements such as "No language has *x*, therefore no state has *x*." We'll return to this issue in a moment.

TIME: 28:45 (+1:00)

#30. F. Evidence

So, you might ask, what could be evidence for markedness, then? How do I go about gathering evidence?

Let's look at a specific example: consonant epenthesis.

- The first question is "what is epenthesis?" This question is often *not addressed* when people talk about epenthesis. However, it is **crucial** to ask and answer it.
- For de Lacy (2006), an epenthetic consonant is a consonant in the PhM output that has *no input correspondent*.
- Importantly, this means that the consonant cannot be affected by faithfulness constraints.
- So, only output constraints determine its form.
- If we eliminate the effect of assimilation and dissimilation, this means that epenthetic consonants give us direct insight into markedness relations.

- There are two important points here.
  - One is that we **cannot** just point to a phonological process and say "This is evidence for markedness." We must **prove** that it **should be** evidence for markedness.
  - This proof involves having a theory.
  - You cannot make any arguments outside of a theory.

TIME: 29:45 (+1:00)

#31. F. Evidence

- This point – about theory-dependency of claims – is very important.
- In fact, there are other theories about epenthetic consonants.
- Staroverov (2014) actually says that epenthetic consonants *always* have input correspondents.
- The consequence of Staroverov's theory is that we **cannot** expect epenthetic consonants to show the effects of output constraints – they are always deeply influenced by faithfulness constraints.
- So, in Staroverov's theory, epenthetic consonants are **not** clear evidence for markedness relations!
- We will assume de Lacy's theory. However, the point here is that we cannot point to a process and say "This is evidence for markedness." We can only say "In this particular theory, the theory states that this process is evidence for markedness."

TIME: 30:30 (+0:45)

#32. F. Evidence

- OK, so if epenthetic consonants have no input correspondent, then what do we need to demonstrate?
- For an epenthetic consonant, we have to show that the consonant is not present in the input, *and* – in this case – that it is not influenced by assimilation or dissimilation.
- How hard is it to do this?

- I wrote a paper with John Kingston about how difficult it is to do this. It's actually really difficult.
- You need morpho-phonological alternations. Here's a sample in this box. You need all of these forms, and *then* prove that it's productive. That's demanding, but it's what the theory requires.
- Often, morphemes, or suppletive allomorphs, *look* like epenthesis, but are not.
- Anyway, the point is that there is a complicated process to prove that there is consonant epenthesis. It is rare that people actually prove that they have epenthesis.

TIME: 31:15 (+0:45)

#33. F. no [k]

- But the big question, and the one most relevant for markedness, is "How do we prove that something doesn't happen?" Or, more formally, "How do we prove that a particular set of states with property *p* do not exist?"
- Let's look at a specific claim: that there is no epenthetic [k].
- The first issue is commonality. How likely is that no epenthetic [k] is an accidental gap?
- Well, [k] itself is common. In 97% of languages from UPSID.
- Epenthesis is probably somewhat common. I don't know *how* common, but I guess it's one of the more common phonological operations.
- And what *can* be epenthetic? Possibly only a few things.
- So, what are the chances of there being epenthetic [k]? Well, that can be determined by estimating how many speakers have epenthesis *and* lexical [k] (by multiplying the probability of [k] with the probability of epenthesis), and then – roughly – dividing by the number of potential epenthetic elements.
- In this case, the number is probably relatively large.

- Apart from commonality, we need to ask "Is [k] learnable?" Specifically, given particular input, is the learning process capable of converging on a state with epenthetic [k]?

    - The answer is yes in that we can model its learning in extant theories of learning.

TIME: 32:15 (+1:00)

#34. F. no [k]

We also have two other relevant questions.

Is epenthetic [k] **actuable**? Is there a plausible circumstance where epenthetic [k] could emerge?

De Lacy & Kingston (2013) argue that yes, there is. It could arrive through misperception at vocalic boundaries.

There is also a lot of diachronic evidence that *t turns into [k]. If there are epenthetic [t]s, then whatever motivates t>k could also favor epenthetic [k].

Finally, we should ask "Is epenthetic [k] **transmissible**?"

Some things are hard to transmit. For example, it's easy for a language to end up with word-initial glottal stops through misinterpretation of articulation. However, it's *difficult* for word-initial glottal stops to survive as contrastive elements – see many Polynesian languages.

In contrast, it's really hard to actuate clicks, but really easy to transmit them.

So, what about epenthetic [k]? We know that lexical [k] is easy to transmit. We know that epenthesis is fairly easy to transmit. Therefore, epenthetic [k] must be easy to transmit.

TIME: 33:15 (+1:00)

#35. F. no [k]

Having asked these questions, we can see that the claim that epenthetic [k] does not occur is not easy to ascribe to coincidence.

However, we might expect some further evidence: a situation where we would *expect* epenthetic [k], but it doesn't occur.

For example, Proto-Polynesian had the stops [p t k], and of these, [t] was epenthetic.

In Pre-Hawai'ian, lexical /k/ became glottal stop. [t] remained epenthetic.

However, in modern Hawai'ian, /t/ became [k]. We would have expected [k] to be the new epenthetic element. However, at that moment, apparently, the epenthetic consonant became glottal stop.

> The mystery is why Hawai'ian doesn't have epenthetic [k], then, when it should have inherited it. De Lacy (2006)'s theory has an explanation: there is no PhM state with epenthetic [k].

TIME: 34:00 (0:45)

#36. G. Current Issues

I have talked about what markedness is, and how we might go about determining whether we have markedness. It's difficult.

A number of people have objected to markedness, both in theories, and in evidence. Instead of talking about all of the objections, which would take a long time, I'm going to identify some false arguments in a general way.

I see lots of mistakes when people talk about markedness. I'm going to identify a few here.

- One common problem is confusing taxonomic and cognitive markedness. This is very common. A random example: I just read a paper written a few months ago which says that markedness is a "basic state", and then lists taxonomic markedness properties.
- A similar problem is listing statements that make sense in taxonomic terms, but not cognitive ones. For a number of examples, see Blevins (2004)'s *Evolutionary Phonology*.
- A further problem is the belief that somehow 'markedness' is pan-theoretical, or atheoretical. It is not. You have to specify which theory you're talking about.
    - For example, I have seen statements like "Hawai'ian lacks [t], so [t] is not unmarked". But in which **theory**? Not in de Lacy (2006)'s. Yes in some others.
- I have also seen overgeneralization of results.
    - For example, Mielke (2004) states that processes that refer to only part of a class, or a disjoint class, are problematic for innatist theories where all place features spread as a constituent.
    - However, any PhM theory that has blocking, opacity, or dual processes – which might be all of them – can do both partial class and disjoint class behavior. So, here, we have an example of where a result for very specific theories doesn't generalize to other theories.

TIME: 35:30 (+1:30)

#37. G. Current Issues

- Further errors involve making assumptions about philosophical underpinnings.
    - I see this a lot in discussions of 'crazy rules'.
    - For a formalist, a crazy rule is a rule that ***cannot be formulated.***
    - For a functionalist, a crazy rule is a rule that **has no functional motivation**.
    - So, exactly what is crazy depends on the theory, and the philosophy behind the theory.
    - So, if someone asserts that there is a 'crazy rule', it's important to know **in which theory**. I rarely see this kind of discussion. Usually, it's just "this rule is clearly crazy", without any demonstration of *which theories it's crazy in*.

- Finally, I also see a lot of bad phonology. For example, I've seen lots of citations of cases as 'epenthesis' without *proving that it's epenthesis*, in even the most rudimentary way. For any case of epenthesis, we need to see proof of the underlying forms involved in the alternations, and proof of productivity. Without these, we have no cases. For excellent examples of how to do it right, see Staroverov (2014).

TIME: 36:15 (+0:45)

#38. G. Current Issues

Perhaps the biggest issue I see in discussions of markedness, and maybe phonology generally, is the lack of careful consideration of the theory.

- Often, we see assumptions that a particular process is evidence for markedness.
    - But in which theory? Theories are different!
- I think this is part of a bigger problem: Atheoretism. Or the idea that science can proceed in a theory-neutral way. This idea was rejected long ago in the philosophy of science (Medawar 1969), but it still seems to be present in phonological work.

- If I was going to point out what I think are **real** and **serious** problems with work on markedness, it would be what I think are **real and serious** problems with *all phonological research*.

    - These problems involve methodology and evidence.

TIME: 37:00 (+0:45)

#39. Methodology

Whenever we gather evidence for our theories, whether it's through fieldwork, or grammars, or spectrograms, or palatography, or MRIs, and so on, we use procedures – a method.

For any method, we **must** answer three questions:

Is it reliable? If we apply the method to the same data over and over, do we get the same result?

Is it precise and accurate?

- We want our methods to be both precise and accurate.
    - We want our results to be clustered around the truth, and within a small range.

- So, are the methods we use as phonologists reliable, accurate, and precise?

TIME: 37:45 (+0:45)

#40. Methodology Reliable Gujarati

- The overwhelming amount of evidence for phonological theories comes from grammars.
- So, is the 'Using grammars method' reliable?
- Recently, Shu-hao Shih (2018) reported on Gujarati stress. He created this useful table.
    - It's a list of all the grammars that have talked about stress in Gujarati.
    - Turner (1921) says that stress always falls on the penultimate syllable. Every vowel is treated the same for stress.
    - In contrast, Cardona (1965) says that it is first attracted to [a] (for 3$ words). If there is no [a], then to a mid vowel or [u], then to [i], then finally to schwa.
    - The other people say other things.
- Given these grammars, can we describe Gujarati stress? No.
- So, are grammars reliable? Clearly not. They all say different things.

By the way, the importance of this finding is that it destroys the main evidence for sonority-driven stress. Sonority-driven stress is the main evidence for markedness conflation.
So, there's no very little evidence for markedness conflation now.

TIME: 39:00 (+1:15)

#41. Methodology VOT

There's a really good example of accuracy and precision for methods used to measure Voice Onset Time. VOT is a useful measure for lots of things, including aspiration and stress.

Francis et al. compared 6 methods of measuring VOT. They assumed that a laryngograph gave perfect measurements. In comparison to the laryngograph, you can see that every other method had a measure of inaccuracy and imprecision.

For example, the f0 method was inaccurate by 2-5ms, but had low imprecision.

The F2 method for aspirated stops was inaccurate by 12-20ms, with a higher imprecision (around 5ms).

You can immediately see here that which method of VOT measurement you choose will influence your results. If you use the F2 method for aspirated stops, you know that your measurements are around 15ms too long, and only within 5ms of the actual value. So, if you measure [p^h] as having 80ms, it's probably between 62.5-67.5ms long.

What's remarkable about this is that **there are so few examples** of phonological - and phonetic! – methods that have been examined for precision and accuracy.

The implications are profound in phonology. If you see someone say "Stressed syllables are 5ms longer than unstressed ones", then you must ask "How accurate and precise is your method?" If it's not precise to within 5ms, then the meaning of their results is indeterminate.

I find it deeply concerning that we do not know what the reliability, accuracy, and precision of most phonological methods are. We assume that grammars are useful, and it's common to use F2 to find the beginning of vowels. But are these method s valid?

TIME: 40:30 (+1:30)

#42. Returning to epenthesis

Now, you might think "OK, but these are surely just problems for more complicated things like stress and VOT. What about simple, obvious things like epenthesis?"

Well, unfortunately, that's not true. I see very little careful work that demonstrates that putative epenthesis is actual epenthesis.

In fact, Staroverov (2015) looked at what – from the grammar – looked like a robust case of epenthesis in Buryat, and found that it wasn't productive.

Staroverov (2014) showed that a case that often gets cited as [t] epenthesis – Apurucayali – isn't actually epenthesis.

And the bigger question is "What's the reliability, precision, and accuracy of epenthesis methodology?"

The answer is: We don't know.

TIME: 41:30 (+1:00)

#43. Conclusions 1

So, let's conclude.

- If someone says "markedness"
    - Are they talking about taxonomic or cognitive markedness?
    - Do they know?
- If they are talking about cognition, then
    - What's the theory?
    - We have a right to know! We *must* know, otherwise we don't know what they're talking about.
    - There's no evidence without a theory.
- If someone says "Here's evidence for or against markedness", they need to demonstrate that
    - The theory claims that the PhM is responsible for that evidence
    - And the theory doesn't assign the responsibility to other modules and mechanisms.

TIME: 42:15 (+0:45)

#44. Conclusions 2

More generally, if someone says “Here’s some evidence”, we need to ask “Is your method valid?” Is it reliable? Precise? Accurate? How do you know?

So, what should we all do now.

We have a lot of theories. Some of them are very well defined and precise.
So now let’s study methodological validity.
When we know about our methods, we can gather evidence.
Then we’ll know whether our markedness theories are right.

TIME: 43:00 (+0:45)

#45. Final Slide

- Thank you. Now let’s discuss.

Show the final slide.

END