Uncertainty as an Organizing Principle of Phonological Systems

Elizabeth Hume (Univ. of Canterbury)
(beth.hume@canterbury.ac.nz)
Andy Wedel (Univ. of Arizona)
(wedel@u.arizona.edu)
Collaborator: Kathleen Currie Hall (Univ. of British Columbia (kathleen.hall@ubc.ca)

13th Biennial Language and Society Conference
University of Auckland

1

Enhancement Contexts

- Certain contexts tend to attract enhanced phonological patterns
  - sounds with better acoustic/auditory cues
  - greater number of phoneme categories
- Word level: word-initial position
- Sub-lexical level:
  - prosodically prominent positions (e.g. stressed syllables, long vowels)
  - segmentally prominent positions (e.g. pre-vocalic consonant, **C**V)

2

Reduction Contexts

- By contrast, reduced phonological patterns tend to occur in the complement contexts
  - sounds with weaker acoustic/auditory cues
  - fewer number of phoneme categories
- Word level: word-final position
- Sub-lexical level:
  - prosodically non-prominent positions (e.g. unstressed syllables, short vowels)
  - segmentally non-prominent positions (e.g. nasal consonant before another consonant, V**N**CV)

3

Previous explanations

Several interrelated explanations have been proposed:

- processing considerations (Cutler (various), Beckman 1997, Smith 2004): earlier parts of words are important for speech processing; how do other "prominent" positions fit in when not word-initial?
- perceptual salience (e.g. Steriade 1997): optimize perceptual distinctiveness of sounds; how do word effects fit in?
- functional needs (e.g. Boersma 1997): interacting functional constraints (e.g. ease of articulation, perceptual distinctiveness) operate on features to generate observed outputs; contextual prominence?

4

Our goal

- To provide a unified account of the patterns of enhancement and reduction that occur at sub-lexical, lexical and higher levels.

- Leading idea: Language is conceived of a system of optimizing communication through efficiently **resolving uncertainty** in the mapping between signal and meaning.
- Patterns of enhancement and reduction can be seen as responses to the need to reduce uncertainty associated with predicting the outcome of a message.

5

A starting point: Communication

- Language involves the transmission of information.
- Speakers encode the message onto a signal; listeners decode the signal back onto a message.
- Linguistic signals are structured in terms of nested and ordered categories
  - Features → phonemes → larger sublexical units → words → phrases → …
- For the listener, the process of decoding a signal can be thought of as *the problem of mapping a noisy signal onto the intended categories*.

6

Why is this a 'problem'?

Noise in the channel (in the environment, in production, in perception...) and the possibility of more than one outcome introduces uncertainty in the signal-category mapping.

→ Consequently, successful information transmission is dependent on the resolution of uncertainty in the mapping between the signal and the message.

7

Resolution of uncertainty

- It is well-established from psycholinguistic and phonetic research that we use two general types of information to resolve uncertainty in signal-category mappings:
  - top-down information from our knowledge of the system, e.g.,
    - possible lexical categories in the sentence context
    - possible phonemic contrasts in a given segmental context
  - bottom-up information from the signal itself, e.g. phonetic cues
- Bayes' Rule provides a way to conceptualize this and unify these two types of information.

8

Bayes' Rule and resolution of uncertainty in signal-message mapping

The probability of the category C in the context

The probability that the signal S is produced, given the category C

The probability of the category C given the signal S in the context

$$p(C|S,ctxt) = \frac{p(C,ctxt)p(S|C)}{\sum_i p(C_i,ctxt)p(S|C_i)}$$

S = the speech signal
C = a category
ctxt = contextual information

9

Anything that increases the value of the numerator decreases the uncertainty in the mapping from a signal S to a category C.

The probability of the category C in the context

The probability that the signal S is produced, given the category C

The probability of the category C given the signal S in the context

$$p(C|S,ctxt) = \frac{p(C,ctxt)p(S|C)}{\sum_i p(C_i,ctxt)p(S|C_i)}$$

S = the speech signal
C = a category
ctxt = contextual information

10

What types of manipulations decrease uncertainty in the mapping from a signal to a category?

- Bottom-up via the signal: increase phonetic distinctiveness
  - Increases p(S|C), the probability of the signal given the category.
- Top-down via the system: decrease the number of categories expected in the context
  - Increases p(C, ctxt), the probability of the category in context

11

The Resource Issue

- Manipulations are constrained by the fixed number of resources available given the system.
  - e.g. We don't have endless amounts of effort to improve the signal. For the speaker, effort relates in part to articulatory effort. For the listener, it relates to attention.

→ Optimal communication systems deploy resources strategically, in ways that are most efficient for communicating the message.

12

Manipulating Redundancy

- In *A Mathematical Theory of Communication* (1948), Claude Shannon proved mathematically that communicating a message through a noisy channel can approximate the ideal when the substance used to encode the message is manipulated in specific ways.
- Redundancy in the signal is increased in contexts of higher uncertainty, and decreased in contexts of lower uncertainty.
- Note that redundancy does not imply wastefulness. Rather, redundancy means that there are multiple cues to the same message.

13

Language as a communication system

- Given that language is a communication system, we expect it to function like communication systems more generally.
- A plethora of laboratory findings are consistent with this expectation.
  - language users pay more attention to elements in prominent positions (e.g., Cutler, Hawkings & Gilligan 1985, MacWhinney 2005, Jurafsky & Martin 2008**)**, tend to enhance or hyperarticulate elements in such positions (e.g., Lindblom 1990), and reduce and delete elements in non-prominent positions (e.g., Lavoie 2002).

14

Modulating uncertainty in usage

- Psycholinguistic evidence that details of speech are modulated to distribute uncertainty evenly across the signal:
  - Word choice and uncertainty at the sentence level (Levy & Jaeger 2007)
  - Phonetic detail and uncertainty at the levels of word and syllable (Van Son & Pols 2003, Aylett & Turk 2004)

15

Prediction

Under the assumption that patterns in grammar originate as patterns of usage, phonological patterns should also show the effects of modifying the message in a way that responds to the degree of uncertainty associated it.

16

Resolving Uncertainty: Predictions for phonology

Two responses to resolving uncertainty in a noisy system:

1. increase the predictiveness of the signal by adding more or better cues (low level response)
2. decrease the number of competing outcomes (higher level response)

When uncertainty is low:

1. decrease predictiveness of cues
2. increase the number of possible outcomes

- Recall that these responses relate back to the two terms in Bayes' rule.

17

Structure of the remainder of the talk

Present evidence consistent with the notion that many phonological patterns optimize uncertainty levels.

18

Maximizing Information

There are different positions within a word that could be modified in order to reduce uncertainty.

Most efficient

↓

Concentrate information in places where it is going to be most effective in resolving uncertainty.

19

Word Processing

- Psycholinguistic evidence shows that parsing words is incremental; we start making decisions about what words are before the end of the word.
- Lexical access is generally achieved on the basis of the initial part of the word (Cutler et al. 1985, Marslen-Wilson 1989, Marslen-Wilson & Zwitserlood 1989).
- This means that information occurring earlier in the parse will reduce uncertainty more than information occurring latter.

20

Predictions for phonological patterns at the word-level

All else being equal:

1. The inventory of available phonemic contrasts should be higher earlier in the word.
2. Better phonetic cues should be favored earlier in the word.

21

Category differences: Root-initial vs. non-initial
(Beckman 1997)

language	*inventory includes*	*initial σ*	*non-initial σ*
Tuva (Turkic) (Krueger 1977)	plain & glottalised vowels	plain & glottalised vowels	no glottalised vowels
!Xóõ (Bushman) (Traill 1985)	click & non-click consonants	click & non-click consonants	no clicks
Tamil (Christdas 1988, Bosch & Wiltshire 1992)	high, mid & low vowels round & unround vowels	high, mid & low vowels round & unround vowels	no mid vowels no round vowels
Turkic family (Comrie 1981, Kaun 1995)	round & unround vowels	round & unround vowels	round vowels only after a round vowel in the initial syllable
Shona (Bantu) (Fortune 1955)	high, mid & low vowels	high, mid & low vowels	mid only after mid in the initial syllable
Dhangar-Kurux (Dravidian) (Gordon 1976)	oral & nasal vowels long & short vowels	oral & nasal vowels long & short vowels	no nasal vowels no long vowels
Shilluk (Nilotic) (Gilley 1992)	plain, palatalised & labialised consonants	plain, palatalised & labialised consonants	no palatalised or labialised consonants
Malayalam (Dravidian) (Wiltshire 1992)	labial, dorsal & a variety of coronal consonants	independent place of articulation in coda position	place of articulation in coda must be shared by following onset
Bashkir (Turkic) (Poppe 1964)	high & non-high vowels front and back vowels	high & non-high vowels front and back vowels	no high vowels front/back only in agreement with σ_1 vowel
Damin (Lardil secret language) (Hale 1973)[5]	Lardil segments, plus nasalised clicks, bilabial & velar ejectives, ingressive lateral fricative	Lardil segments, plus nasalised clicks, bilabial & velar ejectives, ingressive lateral fricative	no clicks, ejectives or lateral fricative

22

Word-level phonetic cue differences

Fortition in word-initial position:

Luganda (Cole 1967): gemination in word-initial position

Class 5 (SG)	Class 6 (PL)	
ggi	ma-gi	'egg'
ddaala	ma-daala	'ladder'
zzike	ma-zike	'chimpanzee'
jjuba	ma-yuba	'dove'
gg^{w}aanga	ma-waanga	'nation'
ddaanga	ma-laanga	'lily'

Lenition in word-final position:

American English: debuccalization of /t/ word-finally, e.g. cat [kæt] ~ [kæʔ]

"of all phonological positions word-initial is the most resistant to deletion." (Harris 2009)

23

Predictions for phonological patterns at sub-lexical level

All else being equal:

1. The inventory of phonemic contrasts should be greater in prosodically or segmentally prominent positions, e.g.
   - prosodic prominence, stress, length
   - segmental context: in contexts where cues are salient (e.g. contexts where cues are not masked, are dissimilar from neighbouring sounds)
2. Phonetic cues should be enhanced in prosodically or segmentally prominent positions

24

Differences in category number

- Stressed syllables vs. unstressed syllables
  - more vowel categories in stressed position (e.g. English)
- Onset position vs. coda postion
  - Cypriot Greek: Consonant fortition in onset position after a consonant (excluding nasals and liquids) (Newton 1972)
  - Porteño Spanish: Glides /y, w/ are pronounced as homorganic obstruent fricatives in syllable-initial position (Lozano 1979).

25

Differences in cue quality

- Segmentally and/or prosodically (non-)prominent positions: reduction (e.g. assimilation, lenition) commonly targets sounds in contexts where the cues are weak

E.g. syntagmatically similar sounds

/nl/ reduction (Seo 2004)

➔ [ll] in Klamath, Ponapean, Toba Batak, Moroccan Arabic, Leti, Korean, Uyghur

➔ [nn] in Tatar, Yakut

➔ [n] in Zoque

26

Getting from phonetic variation in usage to phonological patterns in grammar

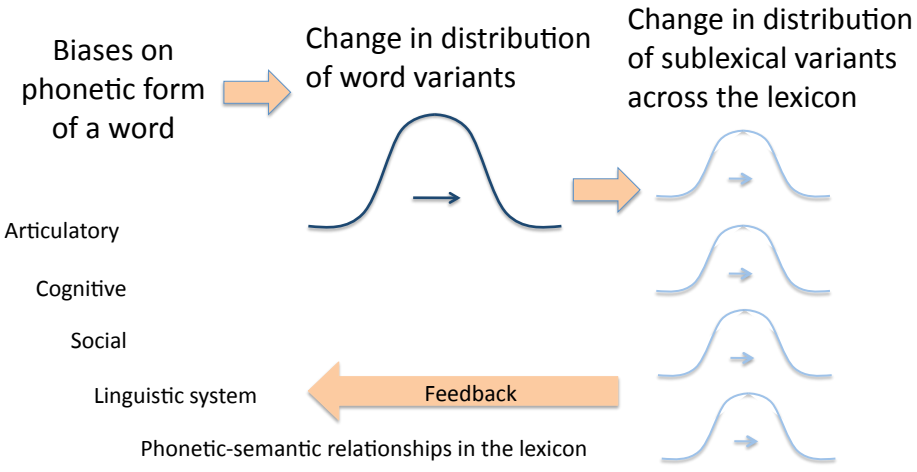
27

'Variationist/Usage Based/Evolutionary' models propose a causal link between utterance level biases and the development of abstract patterns

(see work by Baudoin de Courtenay, Ohala, Lindblom, Blevins, Pierrehumbert, Bybee and many others)

- Based on evidence that:
  - Detailed variants of the same word coexist and compete in the mental lexicon;
  - Variants arise under the influence of biases in articulation, perception, transmission;
  - Rule-like behavior arises from generalization over existing, remembered words.
    - Existing patterns contribute biases to language production and perception, creating feedback.

28



Theoretical/Linguistic evidence:

e.g., Baudouin de Courtenay 1895, Ohala 1989, Lindblom 1990, Blevins 2004, Bybee 2001, Christiansen & Chater 2008

Hay and Maclagan 2012, reviewed in Ernestus 2011

Wang 1969, Bybee 2002, Phillips 2006

29

Summary

- Language is a system for the transmission of information.
- The form of many phonological patterns is consistent with the hypothesis that grammars evolve to distribute information efficiently across the signal.
- Psycholinguistic evidence for corresponding biases in speech production is consistent with evolutionary/exemplar models for grammatical pattern development.

30

System evolution

- The evidence suggests that there's a causal link between the variation that people produce and the patterns observed in phonological systems.
- Both phonetic and phonological patterns provide evidence consistent with the view that the signal is manipulated to resolve message uncertainty.
- This suggests that language systems evolve to concentrate information in places where they're going to do more work.

31

The Primacy of Words

- The emphasis in this approach is on the message (word).
- We resolve uncertainty with regards to a certain category level and the evidence is consistent with the hypothesis that the most important level is the word and not the phoneme.
- All of the phenomena discussed are consistent with the word-level analysis while only a subset are consistent with the phoneme level analysis,
  - e.g. why delete a consonant at the end of the word? It's because it does little work in identifying the **word**, not what the phoneme is.

32

Conclusion

When language is viewed from the perspective of a system designed for the transmission of information, enhancement and reduction can be seen as mechanisms aimed at resolving uncertainty regarding a message.

33

References

Cole, D. T. (1967). *Some features of Ganda linguistic structure.* Johannesburg: Witwatersrand University Press.

Cutler, A., J.A. Hawkins & G. Gilligan (1985) The suffixing preference: a processing explanation. *Linguistics* 23. 723-58.

Lozano, M. (1979). Stop and spirant alternations: fortition and spirantization processes in Spanish phonology. Indiana University Linguistics Club.

Marslen-Wilson, W.D. (1989) Access and integration: Projecting sound onto meaning. In W.D. Marslen-Wilson (ed.), *Lexical Representation and Process*. Cambridge, MA: MIT Press. 3-24.

Marslen-Wilson, W.D., & P. Zwitserlood (1989). Accessing Spoken Words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance.*

Newton, Brian (1972). *Cypriot Greek: its phonology and inflections.* The Hague: Mouton.

34

References (cont.)

Shannon, Claude. 1948. A Mathematical Model of Communication.

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Unpublished ms., University of California, Los Angeles.

35