Corpus data vs. experiments in English phonotactics

Michael Hammond

U. of Arizona

Outline

Collaborators

- Jeff Berry
- Jordan Brewer
- Lynnika Butler
- Jason Ginsburg
- Ben Tucker

What does phonology say about syllable onsets?

What does phonology say about syllable onsets?

- Onset substring Big onsets are more marked than smaller onsets.

What does phonology say about syllable onsets?

- Onset substring Big onsets are more marked than smaller onsets. If a language has an onset $C_1^+ C_2^+$, then it will also have onset C_1^+ and onset C_2^+.

What does phonology say about syllable onsets?

- Onset substring Big onsets are more marked than smaller onsets. If a language has an onset $C_1^+ C_2^+$, then it will also have onset C_1^+ and onset C_2^+.
- Sonority sequencing Onsets with more than one consonant increase in sonority.

What does phonology say about syllable onsets?

- Onset substring Big onsets are more marked than smaller onsets. If a language has an onset $C_1^+ C_2^+$, then it will also have onset C_1^+ and onset C_2^+.
- Sonority sequencing Onsets with more than one consonant increase in sonority. If a language has an onset $\ldots C_1 \ldots C_2 \ldots$, then C_1 is less sonorous than C_2.

Are these true in English?

Are these true in English?

What does it mean to be true?

Are these true in English?

What does it mean to be true?

- Grammar

Are these true in English?

What does it mean to be true?

- Grammar
- Frequency

Are these true in English?

What does it mean to be true?

- Grammar
- Frequency
- Acquisition

Are these true in English?

What does it mean to be true?

- Grammar
- Frequency
- Acquisition
- Judgments

Are these true in English?

What does it mean to be true?

- Grammar
- Frequency
- Acquisition
- Judgments

Are these all in sync with each other?

Are these true in English?

What does it mean to be true?

- Grammar
- Frequency
- Acquisition
- Judgments

Are these all in sync with each other? If not, what story do we tell?

English consonants

p		t		k
b		d		g
f	θ	s	š	
v	ð	z	ž	
			č	
			ǰ	
m		n		ŋ
		l		
		r		
w			y	h

All of these can be syllable onsets except the ones marked in red (Hammond, 1999); [ŋ] cannot occur at all and [ž] is marginal.

The diphthong [yu]

The diphthong [yu]

- English diphthongs can generally occur with any onset.

The diphthong [yu]

- English diphthongs can generally occur with any onset.
- The sequence [yu] behaves like a diphthong; [y] is part of the vowel, not part of the onset (Davis & Hammond, 1995).

The diphthong [yu]

- English diphthongs can generally occur with any onset.
- The sequence [yu] behaves like a diphthong; [y] is part of the vowel, not part of the onset (Davis & Hammond, 1995).
  - In the sequence CyV, V can only be [u], e.g. in *skew* [skyu], *few* [fyu], etc.

The diphthong [yu]

- English diphthongs can generally occur with any onset.
- The sequence [yu] behaves like a diphthong; [y] is part of the vowel, not part of the onset (Davis & Hammond, 1995).
  - In the sequence CyV, V can only be [u], e.g. in *skew* [skyu], *few* [fyu], etc.
  - Pig Latin

Strom [stram]	⟶	[am-stre]
Gwen [gwɛn]	⟶	[ɛn-gwe]
Beula [byulə]	⟶	[yulə-be]

Complex onsets

Complex onsets

- $\left\{ \begin{array}{c} \text{stop} \\ \text{voiceless fricative} \end{array} \right\} + \left\{ \begin{array}{c} \text{l} \\ \text{r} \\ \text{w} \end{array} \right\}$

  e.g. *play* [ple], *fry* [fray], *queen* [kwin], etc.

Complex onsets

- $\left\{ \begin{array}{c} \text{stop} \\ \text{voiceless fricative} \end{array} \right\} + \left\{ \begin{array}{c} \text{l} \\ \text{r} \\ \text{w} \end{array} \right\}$
  e.g. *play* [ple], *fry* [fray], *queen* [kwin], etc.
- $[\text{s}] + \left\{ \begin{array}{c} \text{nasal} \\ \text{voiceless stop} \\ \text{(voiceless fricative)} \end{array} \right\}$
  e.g. *snow* [sno], *spot* [spat], *sphere* [sfir], etc.

Complex onsets

- $\left\{ \begin{array}{c} \text{stop} \\ \text{voiceless fricative} \end{array} \right\} + \left\{ \begin{array}{c} \text{l} \\ \text{r} \\ \text{w} \end{array} \right\}$

  e.g. *play* [ple], *fry* [fray], *queen* [kwin], etc.

- $[\text{s}] + \left\{ \begin{array}{c} \text{nasal} \\ \text{voiceless stop} \\ \text{(voiceless fricative)} \end{array} \right\}$

  e.g. *snow* [sno], *spot* [spat], *sphere* [sfir], etc.

- $[\text{s}] + \left\{ \begin{array}{c} \text{voiceless stop} \\ \text{(voiceless fricative)} \end{array} \right\} + \left\{ \begin{array}{c} \text{l} \\ \text{r} \\ \text{w} \end{array} \right\}$

  e.g. *splash* [splæš], *spree* [spri], *squash* [skwɔš], etc.

Grammatical generalizations

Grammatical generalizations

Remember: neither generalization is a biconditional.

- Onset substring
- Sonority sequencing

Grammatical generalizations

Remember: neither generalization is a biconditional.

- Onset substring True: if $C_1^+ C_2^+$ is an onset, then C_1^+ and C_2^+ are both onsets. (Clements & Keyser, 1983)
- Sonority sequencing

Grammatical generalizations

Remember: neither generalization is a biconditional.

- Onset substring True: if $C_1^+ C_2^+$ is an onset, then C_1^+ and C_2^+ are both onsets. (Clements & Keyser, 1983)
- Sonority sequencing True: if $\ldots C_1 \ldots C_2 \ldots$ is an onset, then C_1 is less sonorous than C_2, except for [s]-clusters, a common exception. (Clements & Keyser, 1983)

Grammatical generalizations

Remember: neither generalization is a biconditional.

- Onset substring True: if $C_1^+ C_2^+$ is an onset, then C_1^+ and C_2^+ are both onsets. (Clements & Keyser, 1983)
- Sonority sequencing True: if $\ldots C_1 \ldots C_2 \ldots$ is an onset, then C_1 is less sonorous than C_2, except for [s]-clusters, a common exception. (Clements & Keyser, 1983)

For example, in [kw] in *queen*, [k] and [w] are independently onsets, e.g. in *keel* [kil] and *we* [wi], and [k] is less sonorous than [w].

Markedness

Markedness

These are really claims about markedness, which makes stronger claims in the domains of frequency, acquisition, and gradient judgments. More marked elements should be:

Markedness

These are really claims about markedness, which makes stronger claims in the domains of frequency, acquisition, and gradient judgments. More marked elements should be:

- less frequent;

Markedness

These are really claims about markedness, which makes stronger claims in the domains of frequency, acquisition, and gradient judgments. More marked elements should be:

- less frequent;
- acquired earlier;

Markedness

These are really claims about markedness, which makes stronger claims in the domains of frequency, acquisition, and gradient judgments. More marked elements should be:

- less frequent;
- acquired earlier;
- judged more well-formed.

Predictions about onsets

Predictions about onsets

- Cluster size
- Sonority

Predictions about onsets

- Cluster size
  - Larger clusters are less frequent;
- Sonority
  - Onsets that violate sonority are less frequent;

Predictions about onsets

- Cluster size
  - Larger clusters are less frequent;
  - larger clusters are acquired later;
- Sonority
  - Onsets that violate sonority are less frequent;
  - onsets that violate sonority are acquired later;

Predictions about onsets

- Cluster size
  - Larger clusters are less frequent;
  - larger clusters are acquired later;
  - larger clusters are less well-formed.
- Sonority
  - Onsets that violate sonority are less frequent;
  - onsets that violate sonority are acquired later;
  - onsets that violate sonority are less well-formed.

Predictions about onsets

- Cluster size
  - Larger clusters are less frequent;
  - larger clusters are acquired later;
  - larger clusters are less well-formed.
- Sonority
  - Onsets that violate sonority are less frequent;
  - onsets that violate sonority are acquired later;
  - onsets that violate sonority are less well-formed.

But there shouldn't be any clusters that violate the Sonority generalization.

How frequent are these clusters in the Brown corpus?

How frequent are these clusters in the Brown corpus?

- Size of clusters

size	tokens
0	302878
1	630175
2	76270
3	3713

($X^2 : p < 2.2e - 16$)

How frequent are these clusters in the Brown corpus?

- Size of clusters

size	tokens
0	302878
1	630175
2	76270
3	3713

$(X^2 : p < 2.2e - 16)$

- Sonority of clusters (2- and 3-consonant clusters)

sonority	tokens
Rising	61571
Falling	18412

$(X^2 : p < 2.2e - 16)$

How frequent are these clusters in the Brown corpus?

- Size of clusters

size	tokens
0	302878
1	630175
2	76270
3	3713

$(X^2 : p < 2.2e - 16)$

- Sonority of clusters (2- and 3-consonant clusters)

sonority	tokens
Rising	61571
Falling	18412

$(X^2 : p < 2.2e - 16)$

- This latter comparison may not be apropos, since [s]-clusters are exceptional.

What happens in acquisition?

What happens in acquisition?

- Levelt, Schiller & Levelt (1999-2000): smaller clusters are acquired before larger clusters.

What happens in acquisition?

- Levelt, Schiller & Levelt (1999-2000): smaller clusters are acquired before larger clusters.
- Meta-analysis of Smit et al. (1993) shows no effect of rising/falling sonority on cluster acquisition.

What happens in acquisition?

- Levelt, Schiller & Levelt (1999-2000): smaller clusters are acquired before larger clusters.
- Meta-analysis of Smit et al. (1993) shows no effect of rising/falling sonority on cluster acquisition. (Again, this comparison may not be apropos, because [s]-clusters are an exception.)

What happens in acquisition?

- Levelt, Schiller & Levelt (1999-2000): smaller clusters are acquired before larger clusters.
- Meta-analysis of Smit et al. (1993) shows no effect of rising/falling sonority on cluster acquisition. (Again, this comparison may not be apropos, because [s]-clusters are an exception.)
- But: D. Ohala (1996) shows that complex clusters are simplified to their least sonorous members.

Up to here

Up to here

- Corpus frequency and acquisition show clear effects of cluster size.

Up to here

- Corpus frequency and acquisition show clear effects of cluster size.
- It's not clear whether or how sonority shows itself in these domains.

Judgments

Judgments

- Generative grammar says that grammaticality judgments are a direct reflection of competence.

Judgments

- Generative grammar says that grammaticality judgments are a direct reflection of competence.
- Are they?

Judgments

- Generative grammar says that grammaticality judgments are a direct reflection of competence.
- Are they?
- Experimental judgment tasks are a way of examining judgments more closely.

Judgments

- Generative grammar says that grammaticality judgments are a direct reflection of competence.
- Are they?
- Experimental judgment tasks are a way of examining judgments more closely.
- Gradient judgments allow us to look even more closely.

Neighborhood density

Neighborhood density

- Nonsense words are judged as well-formed if they sound like lots of actual words (Ohala & Ohala, 1986).

Neighborhood density

- Nonsense words are judged as well-formed if they sound like lots of actual words (Ohala & Ohala, 1986).
- Neighborhood density: how many actual words are "one segment away" from the form in question?

Neighborhood density

- Nonsense words are judged as well-formed if they sound like lots of actual words (Ohala & Ohala, 1986).
- Neighborhood density: how many actual words are "one segment away" from the form in question?
- For example, *blick* [blɪk] has these neighbors:

click [klɪk]	*flick* [flɪk]
lick [lɪk]	*slick* [slɪk]
brick [brɪk]	*black* [blæk]
bleak [blik]	*block* [blak]
blink [blɪŋk]	*bliss* [blɪs]

Phonotactic probability

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).
  - $wf(\text{blɪk}) \approx p(\text{bl}) \times p(\text{ɪk})$

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).
  - $wf(\text{blɪk}) \approx p(\text{bl}) \times p(\text{ɪk})$
  - $wf(\text{blɪk}) \approx p(\mathrm{b}) \times p(\mathrm{l}|\mathrm{b}) \times p(\text{ɪ}|\mathrm{l}) \times p(\mathrm{k}|\text{ɪ})$

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).
  - $\mathit{wf}(\text{blɪk}) \approx p(\text{bl}) \times p(\text{ɪk})$
  - $\mathit{wf}(\text{blɪk}) \approx p(\mathrm{b}) \times p(\mathrm{l}|\mathrm{b}) \times p(\text{ɪ}|\mathrm{l}) \times p(\mathrm{k}|\text{ɪ})$
- This effect is independent of neighborhood density (Bailey & Hahn, 2001).

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).
  - $\textit{wf}(\text{blɪk}) \approx p(\text{bl}) \times p(\text{ɪk})$
  - $\textit{wf}(\text{blɪk}) \approx p(\text{b}) \times p(\text{l}|\text{b}) \times p(\text{ɪ}|\text{l}) \times p(\text{k}|\text{ɪ})$
- This effect is independent of neighborhood density (Bailey & Hahn, 2001).
- Both effects also show up in yes-no tasks (Frisch et al., 2000).

Phonotactic probability

- Nonsense words are judged as well-formed if their parts are frequent (Coleman & Pierrehumbert, 1997).
  - $\textit{wf}(\text{blɪk}) \approx p(\text{bl}) \times p(\text{ɪk})$
  - $\textit{wf}(\text{blɪk}) \approx p(\mathrm{b}) \times p(\mathrm{l}|\mathrm{b}) \times p(\text{ɪ}|\mathrm{l}) \times p(\mathrm{k}|\text{ɪ})$
- This effect is independent of neighborhood density (Bailey & Hahn, 2001).
- Both effects also show up in yes-no tasks (Frisch et al., 2000).
- Both effects show up with auditory or visual presentation (Bailey & Hahn, 2001).

Previous results

Previous results

- If we want to look at the effects of markedness of onsets, we have to factor out the effects of neighborhood density and phonotactic probability.

Previous results

- If we want to look at the effects of markedness of onsets, we have to factor out the effects of neighborhood density and phonotactic probability.
- One way to do that is to carefully select materials so these are all balanced and controlled.

Previous results

- If we want to look at the effects of markedness of onsets, we have to factor out the effects of neighborhood density and phonotactic probability.
- One way to do that is to carefully select materials so these are all balanced and controlled.
- Match items for neighborhood density and phonotactic probability.

Previous results presented here

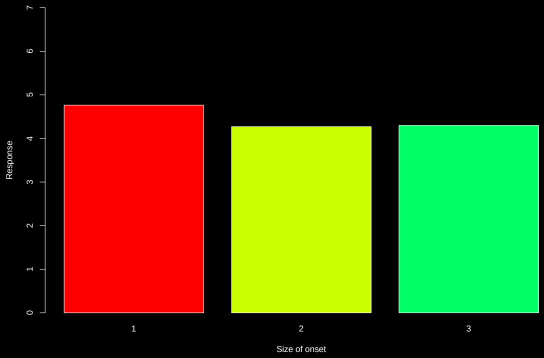
C	CC	CCC
zɪlm	twɪlm	skrɪlm
gayθ	stayθ	strayθ
vork	flork	strork
viʃ	kwiʃ	skriʃ
dʒæntʃ	spæntʃ	stræntʃ
basp	plasp	strasp
rarv	klarv	strarv

Previous results presented here

C	CC	CCC
zɪlm	twɪlm	skrɪlm
gayθ	stayθ	strayθ
vork	flork	strork
viʃ	kwiʃ	skriʃ
dʒæntʃ	spæntʃ	stræntʃ
basp	plasp	strasp
rarv	klarv	strarv

Items matched as closely as possible for neighborhood density and phonotactic probability.

Uh-oh!



But the effect appears to go in the wrong direction.

Many problems

Many problems

- Significant by subjects [$F(2, 19) = 4.31$, $p < .01$], but not by items [F(2,18) = 1.27, n.s.].

Many problems

- Significant by subjects [$F(2,19) = 4.31$, $p < .01$], but not by items [F(2,18) = 1.27, n.s.].
- If we add in neighbors and phonotactic probability, then:
  - there is an effect of neighbors, also in the wrong direction. By subjects: [$F(2,19) = 7.95$; $p < .000$]; By items: [$F(2,18) = 4.87$; $p < .03$]
  - plus several significant interactions of phonotactic probability and neighbors with onset size.

The upshot

The upshot

- A problem: Did we somehow reverse the scale? Is there some unknown factor that overwhelms the factors we're interested in?

The upshot

- A problem: Did we somehow reverse the scale? Is there some unknown factor that overwhelms the factors we're interested in?
- A significant problem: We can't control neighborhood density or phonotactic probability adequately for this investigation.

The upshot

- A problem: Did we somehow reverse the scale? Is there some unknown factor that overwhelms the factors we're interested in?
- A significant problem: We can't control neighborhood density or phonotactic probability adequately for this investigation.
- Teaser: There seems to be a typological effect of onset complexity in there somewhere.

New experiment

New experiment

- How else might we control for the effects of phonotactic probability?

New experiment

- How else might we control for the effects of phonotactic probability? Choose materials that are impossible words of English.

bmluke	dliz	dmloke	fmreap	fnape	fnlope	gnruke
kmroot	kneeb	knliz	ldiz	lmbuke	lmdoke	lmthed
lmube	lmzen	lnfope	lnkiz	lnthem	lntope	lnzape
lshiz	lthiz	lzoog	mlube	mpazz	mruke	mtaz
mthazz	mvupe	mziv	nfape	nkeeb	nzafe	pmazz
pmreeze	rmfeap	rmkoot	rmpeeze	rmthass	rmtofe	rmuke
rmzube	rnguke	rntafe	rntheef	rnvizz	rviss	shliz
thliz	thmazz	thmled	thmrass	thnlem	thnreef	tmaz
tmrofe	tnlope	tnrafe	vmupe	vnrizz	vriss	zloog
zmiv	zmlen	zmrube	znafe	znlape		

New experiment

- How else might we control for the effects of phonotactic probability? Choose materials that are impossible words of English.

bmluke	dliz	dmloke	fmreap	fnape	fnlope	gnruke
kmroot	kneeb	knliz	ldiz	lmbuke	lmdoke	lmthed
lmube	lmzen	lnfope	lnkiz	lnthem	lntope	lnzape
lshiz	lthiz	lzoog	mlube	mpazz	mruke	mtaz
mthazz	mvupe	mziv	nfape	nkeeb	nzafe	pmazz
pmreeze	rmfeap	rmkoot	rmpeeze	rmthass	rmtofe	rmuke
rmzube	rnguke	rntafe	rntheef	rnvizz	rviss	shliz
thliz	thmazz	thmled	thmrass	thnlem	thnreef	tmaz
tmrofe	tnlope	tnrafe	vmupe	vnrizz	vriss	zloog
zmiv	zmlen	zmrube	znafe	znlape		

- Items have no neighbors and have zero phonotactic probability.

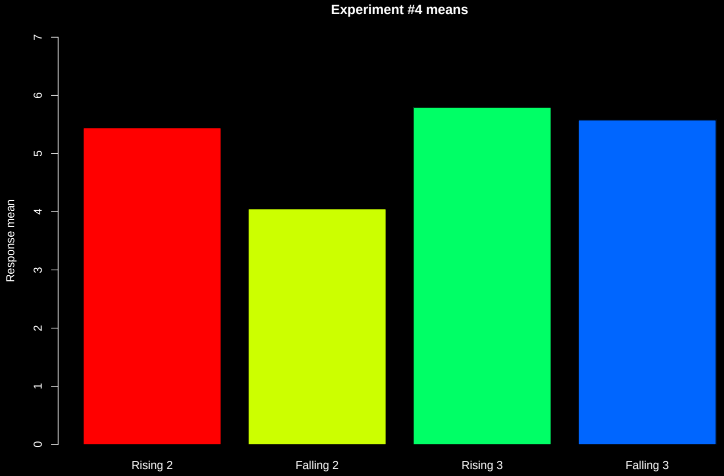
New experiment

- How else might we control for the effects of phonotactic probability? Choose materials that are impossible words of English.

bmluke	dliz	dmloke	fmreap	fnape	fnlope	gnruke
kmroot	kneeb	knliz	ldiz	lmbuke	lmdoke	lmthed
lmube	lmzen	lnfope	lnkiz	lnthem	lntope	lnzape
lshiz	lthiz	lzoog	mlube	mpazz	mruke	mtaz
mthazz	mvupe	mziv	nfape	nkeeb	nzafe	pmazz
pmreeze	rmfeap	rmkoot	rmpeeze	rmthass	rmtofe	rmuke
rmzube	rnguke	rntafe	rntheef	rnvizz	rviss	shliz
thliz	thmazz	thmled	thmrass	thnlem	thnreef	tmaz
tmrofe	tnlope	tnrafe	vmupe	vnrizz	vriss	zloog
zmiv	zmlen	zmrube	znafe	znlape		

- Items have no neighbors and have zero phonotactic probability.
- Factors are size and sonority cline of onset.

Results from new experiment



Statistics

Statistics

- Items with onsets that are bigger are judged as more ill-formed.

  [by subjects: $F(1, 20) = 122.116$, $p < .000$; by items: $F(1, 64) = 37.889$, $p < .000$]

Statistics

- Items with onsets that are bigger are judged as more ill-formed. [by subjects: $F(1, 20) = 122.116$, $p < .000$; by items: $F(1, 64) = 37.889$, $p < .000$]
- Items with onsets that violate the sonority hierarchy are judged as more ill-formed. [by subjects: $F(1, 20) = 75.807$, $p < .000$; by items: $F(1, 64) = 23.521$, $p < .000$]

Statistics

- Items with onsets that are bigger are judged as more ill-formed. [by subjects: $F(1, 20) = 122.116$, $p < .000$; by items: $F(1, 64) = 37.889$, $p < .000$]
- Items with onsets that violate the sonority hierarchy are judged as more ill-formed. [by subjects: $F(1, 20) = 75.807$, $p < .000$; by items: $F(1, 64) = 23.521$, $p < .000$]
- But the factors interact. [by subjects: $F(1, 20) = 47.979$, $p < .000$; by items: $F(1, 64) = 14.886$, $p < .000$]

Up to here

Up to here

- Syllabic markedness plays a role in judgment tasks.

Up to here

- Syllabic markedness plays a role in judgment tasks.
- Do other markedness relations play a role?

Segmental markedness

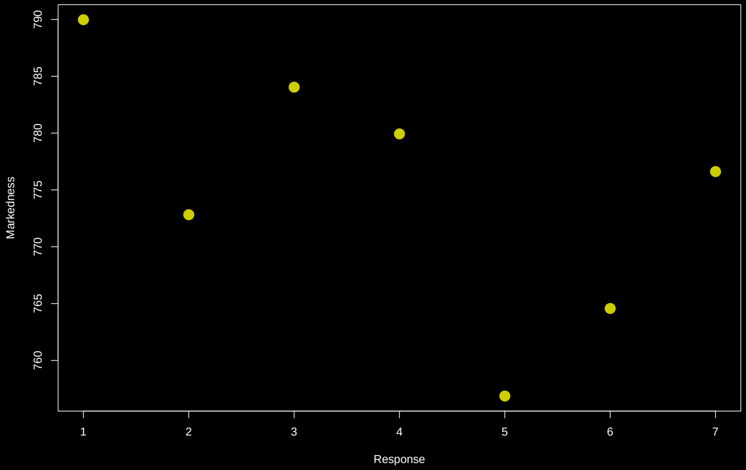
Segmental markedness

- How to calculate segmental markedness: count the number of times each segment occurs in UPSID. Thanks to Natasha Warner!

Segmental markedness

- How to calculate segmental markedness: count the number of times each segment occurs in UPSID. Thanks to Natasha Warner!
- We replicated Bailey & Hahn's visual presentation experiment and found an independent effect of segmental markedness: items containing typologically more marked sounds are judged as less well-formed in the Bailey & Hahn visual presentation replication. [By subjects: phonotactic probability $F(1, 17) = 197.33$, $p < .000$; neighborhood density $F(1, 17) = 46.32$, $p < .000$; markedness $F(1, 17) = 21.96$, $p < .000$] and [By items: phonotactic probability $F(1, 254) = 43.98$, $p < .000$; neighborhood density $F(1, 254) = 10.33$, $p < 0.001$; markedness $F(1, 254) = 4.89$, $p < 0.02$]

The effect of markedness



Desiderata

What properties must a correct model of judgments have?

Desiderata

What properties must a correct model of judgments have?

- a role for phonotactic probability;

Desiderata

What properties must a correct model of judgments have?

- a role for phonotactic probability;
- a role for neighborhood density;

Desiderata

What properties must a correct model of judgments have?

- a role for phonotactic probability;
- a role for neighborhood density;
- a role for markedness;

Toward a model

Toward a model

- Several (cognitive) machines operating in parallel.

Toward a model

- Several (cognitive) machines operating in parallel.
- Each machine produces a numerical result.

Toward a model

- Several (cognitive) machines operating in parallel.
- Each machine produces a numerical result.
- Outputs of machines combined by linear regression.

Toward a model

- Several (cognitive) machines operating in parallel.
- Each machine produces a numerical result.
- Outputs of machines combined by linear regression.
- What machines?

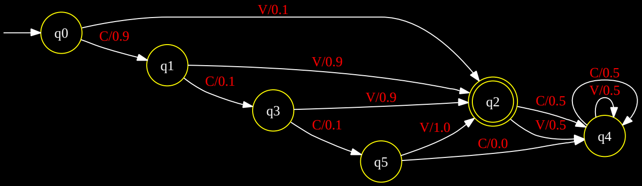
Toward a model

- Several (cognitive) machines operating in parallel.
- Each machine produces a numerical result.
- Outputs of machines combined by linear regression.
- What machines?
  - Weighted Finite automaton (WFSA) for language-specific frequency distributions and for typological patterns. (Initial weights reflect typological predispositions.)

Toward a model

- Several (cognitive) machines operating in parallel.
- Each machine produces a numerical result.
- Outputs of machines combined by linear regression.
- What machines?
  - Weighted Finite automaton (WFSA) for language-specific frequency distributions and for typological patterns. (Initial weights reflect typological predispositions.)
  - Something else for lexical items.

Initial weights for onset size

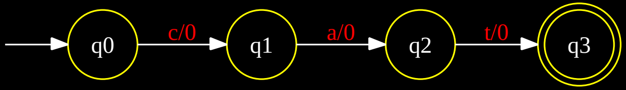


Reweighting by experience

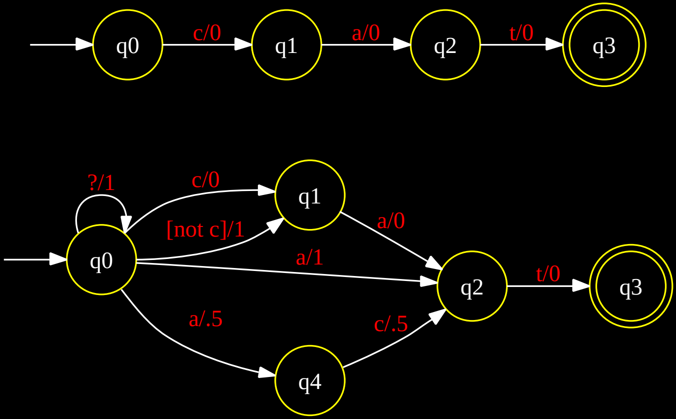
onset	before	after
fV	.009	.09
sfV	.007	.07
sfrV	.005	.05
fnV	.007	.007
fnrV	.005	.005

Weights here are illustrative.

Neighborhood density



Neighborhood density



Why are there different effects?

Why are there different effects?

- Corpora and acquisition don't show obvious effects of sonority cline;

Why are there different effects?

- Corpora and acquisition don't show obvious effects of sonority cline;
- Corpora involve real words;

Why are there different effects?

- Corpora and acquisition don't show obvious effects of sonority cline;
- Corpora involve real words;
- Naturalistic observations of children involve real words;

Why are there different effects?

- Corpora and acquisition don't show obvious effects of sonority cline;
- Corpora involve real words;
- Naturalistic observations of children involve real words;
- Experiments on adults can involve impossible words.

Conclusions

Conclusions

- It can be especially informative to look at impossible words, rather than just possible words.

Conclusions

- It can be especially informative to look at impossible words, rather than just possible words.
- Segmental and syllabic markedness play a role in gradient judgments.

Conclusions

- It can be especially informative to look at impossible words, rather than just possible words.
- Segmental and syllabic markedness play a role in gradient judgments.
- These factors may not play a role in corpus studies or observations of natural acquisition for methodological reasons.

Conclusions

- It can be especially informative to look at impossible words, rather than just possible words.
- Segmental and syllabic markedness play a role in gradient judgments.
- These factors may not play a role in corpus studies or observations of natural acquisition for methodological reasons.
- Typology and neighborhood density can be incorporated into judgments by using WFSAs and setting initial weights accordingly.