Stress, Phrasing, and Auxiliary Contraction in English

ARTO ANTTILA

UC Santa Cruz

March 13, 2015

- English auxiliaries have several alternative realizations:

FULL	REDUCED	CONTRACTED
[wɪɬ]	[wəɬ], [əɬ]	[l]
[hæv]	[həv], [əv]	[v]

- English auxiliaries have several alternative realizations:

FULL	REDUCED	CONTRACTED
[wɪɬ]	[wəɬ], [əɬ]	[l]
[hæv]	[həv], [əv]	[v]

- The FULL vs. CONTRACTED alternation appears to be allomorphy (Kaisse 1983, 1985, Ch. 3)

- English auxiliaries have several alternative realizations:

FULL	REDUCED	CONTRACTED
[wɪɬ]	[wəɬ], [əɬ]	[l]
[hæv]	[həv], [əv]	[v]

- The FULL vs. CONTRACTED alternation seems to be allomorphy. (Kaisse 1983, 1985, Ch. 3)

- The allomorphs are in free variation in some environments.

Environment 1: Contraction is optional

you pay me i’ll do this thing	(The Buckeye Corpus)
You’ll like it in Manitoba	(Zwicky 1970)

Environment 1: Contraction is optional

you pay me i'll do this thing (The Buckeye Corpus)
You'll like it in Manitoba (Zwicky 1970)

Environment 2: Contraction is blocked

*I think, therefore I'm
*Grace and you'll like it in Manitoba

Environment 1: Contraction is optional

you pay me i'll do this thing (The Buckeye Corpus)
You'll like it in Manitoba (Zwicky 1970)

Environment 2: Contraction is blocked

*I think, therefore I'm
*Grace and you'll like it in Manitoba

What is the difference between Environments 1 and 2?

References

1960's: Labov 1969

1970's: Lakoff 1970, King 1970, Zwicky 1970, Baker 1971, Bresnan 1978

1980's: Kaisse 1983, 1985, Sells 1983, Selkirk 1984

1990's: Inkelas and Zec 1993, McElhinny 1993, Pullum 1997, Sadler 1997, Wilder 1997, Krug 1998

2000's: Bender and Sag 2001, Anderson 2008

2010's: MacKenzie 2011, 2012, Bresnan and Spencer 2014, Spencer 2014, Anttila to appear, Barth and Kapatsinski to appear

Proposal 1: Contraction is about stress

Contraction applies to sequences of two unstressed words, e.g., *I will surVIVE* ~ *I'll surVIVE*, and is blocked elsewhere.

Examples

Blocking by lexical stress

- Auxiliaries contract, main verbs don't (*I've got a car* / **I've a car*).

Examples

Blocking by lexical stress

- Auxiliaries contract, main verbs don't (*I've got a car* / **I've a car*).
- The preferred hosts are monosyllabic pronouns (*I'll* / **chiropractors'll*).

Examples

Blocking by lexical stress

- Auxiliaries contract, main verbs don't (*I've got a car* / **I've a car*).
- The preferred hosts are monosyllabic pronouns (*I'll* / **chiropractors'll*).

Blocking by phrasal stress

- Contraction is blocked phrase-finally (*Yes, I WILL* / **Yes, I'LL*).

But how to make this theory work?

We need to be able to determine the

- presence
- absence
- degree

of stress on particular words in particular sentences.

The way forward

Step 1: Adopt an explicit theory of stress.

The way forward

Step 1: Adopt an explicit theory of stress.

Step 2: Formulate an auxiliary hypothesis that connects stress and contraction (= the proposal above).

The way forward

Step 1: Adopt an explicit theory of stress.

Step 2: Formulate an auxiliary hypothesis that connects stress and contraction (= the proposal above).

Given such a theory

- we can derive predictions about the distribution of contraction
- we can use contraction data to test analyses of stress

The way forward

Step 1: Adopt an explicit theory of stress.

Step 2: Formulate an auxiliary hypothesis that connects stress and contraction (= the proposal above).

Given such a theory

- we can derive predictions about the distribution of contraction
- we can use contraction data to test analyses of stress

Advantage: Prominence is hard to hear. Contraction is easier to hear and we can count its application frequency in spoken/written corpora.

A quick review of English stress

A quick review of English stress

1. The Nuclear Stress Rule (NSR):

In a phrase (NP, VP, AP, S), assign stress to the rightmost word bearing lexical stress (= [1 stress]).

A quick review of English stress

1. The Nuclear Stress Rule (NSR):

In a phrase (NP, VP, AP, S), assign stress to the rightmost word bearing lexical stress (= [1 stress]).

2. The Compound Stress Rule (CSR):

In a compound word (N, A, V), skip over the rightmost word and assign stress to the rightmost word bearing lexical stress (= [1 stress]); if there is none try again without skipping.

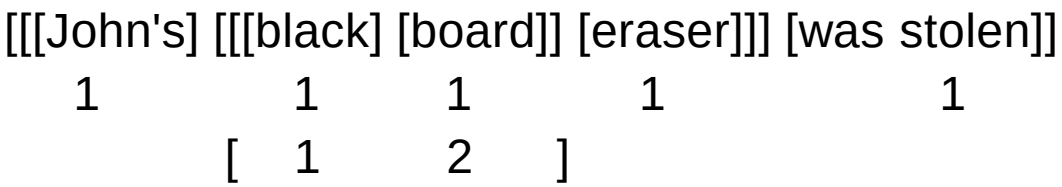
The cycle

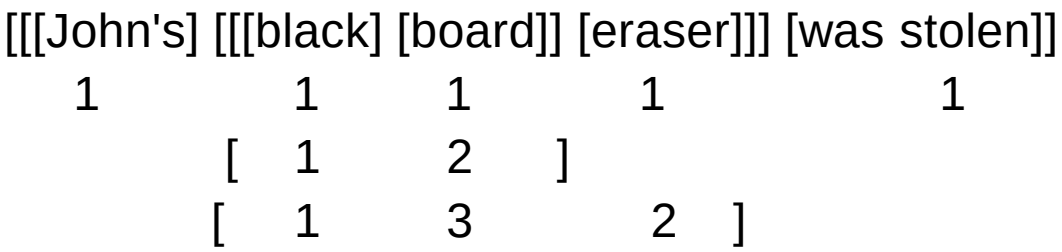
The CSR and the NSR apply cyclically, starting from the innermost brackets, assigning [1 stress] and reducing stress elsewhere by one (stress subordination).

Example

```
                                                              x
                       x                                      x
        x              x                                      x
        x              x                    x                 x
        x              x         x          x                 x
[ [ [ John's ] [ [ [ black ] [ board] ] [ eraser] ] ] [ was stolen ] ]
        3              2         5          4                 1
```

[[[John's]	[[[black]	[board]]	[eraser]]]	[was stolen]]
1	1	1	1	1


[[[John's] [[[black] [board]] [eraser]]] [was stolen]]
1 1 1 1 1
[1 2]


[[[John's] [[[black] [board]] [eraser]]] [was stolen]]
1 1 1 1 1
[1 2]
[1 3 2]

[[[John's] [[[black] [board]] [eraser]]] [was stolen]]

1 1 1 1 1

[1 2]

[1 3 2]

[2 1 4 3]

```
[[[John's] [[[black] [board]] [eraser]]] [was stolen]]
    1          1       1         1              1
               [   1       2     ]
               [   1       3         2   ]
 [  2          1       4         3     ]
[   3          2       5         4              1    ]
```

Problems

Lexical stress

Are all monosyllabic function words, e.g., *will, shall, who, you, have, is, was, it,* etc. lexically unstressed to the same degree?

(Ladd 1980, O'Shaughnessy and Allen 1983, Altenberg 1987, Baart 1987, Hirschberg 1993, Shih 2014)

Problems

Lexical stress

Are all monosyllabic function words, e.g., *will, shall, who, you, have, is, was, it,* etc. lexically unstressed to the same degree?

(Ladd 1980, O'Shaughnessy and Allen 1983, Altenberg 1987, Baart 1987, Hirschberg 1993, Shih 2014)

Phrasal stress

The Nuclear Stress Rule is a good first approximation of default phrasal stress, but in actual sentences we find a lot of variation.

Proposal 2: Lexical stress allows gradation

- Some words are more stressable than others.

Proposal 2: Lexical stress allows gradation

- Some words are more stressable than others.
- Let us call the stressability of a word its STRENGTH.

Proposal 2: Lexical stress allows gradation

- Some words are more stressable than others.

- Let us call the stressability of a word its STRENGTH.

STRENGTH	EXAMPLES	WORD CLASS
1	*it*	weak pronouns
2	*you, that, is, am,* $have_{AUX}$	strong pronouns, finite auxiliaries
3	*could, will, how*	modals, WH-words
4	*stolen, John,* $have_{LEX}$	open class words

Lexical stress as a stringency hierarchy

Lexical stress:

Assign a violation for every lexical item of strength *n* with phrasal stress.

(a)	*STRESS/1	No phrasal stress on Class 1.
(b)	*STRESS/12	No phrasal stress on Classes 1 or 2.
(c)	*STRESS/123	No phrasal stress on Classes 1 or 2 or 3.
(d)	*STRESS/1234	No phrasal stress on Classes 1 or 2 or 3 or 4.

Proposal 3: The NSR as a gradient constraint

Phrasal stress:

The Nuclear Stress Constraint (NSC): Assign a violation for each word between phrasal stress and the right edge of the phrase.

tic tac toe	NSC
(a) ☞ tic tac TOE	
(b) tic TAC toe	1!
(c) TIC tac toe	2!

Other constraints

*WORD	Assign a violation for every word.
FAITH	No contraction.
FAITH/NSC	No contraction under phrasal stress.

The core of the analysis

- Phrasal stress (NSC) goes as far right as possible.
- Markedness (*STRESS/n) prefers stress on strong words.

Contraction is blocked phrase-finally

it will 1 3	FAITH/NSC	NSC	*S/1	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ it WILL					1	1		2
(b) IT will		1	1	1	1	1		2
(c) IT'LL	*!		1	1	2	2	1	1

Contraction is possible if a stronger word follows

she will go 2 3 4	NSC	*S/1	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ she will GO					1		3
(b) ☞ she'll GO					1	1	2
(c) she WILL go	1			1	1		3
(d) SHE will go	2		1	1	1		3

Variation: FAITH >> *WORD (= no contraction)

she will go 2 3 4	*S/1234	FAITH	*WORD
(a) ☞ she will GO	1		3
(b) she'll GO	1	1 **W**	2 **L**

Variation: *WORD >> FAITH (= contraction)

she will go 2 3 4	*S/1234	*WORD	FAITH
(a) she will GO	1	3 **W**	**L**
(b) ☞ she'll GO	1	2	1

The theory of variation

- An individual's competence is not a total order, but a PARTIAL ORDER (see e.g., Kiparsky 1993, Anttila 1997, Anttila and Cho 1998, Zamma 2013, Djalali 2013).

- Variation arises in performance as the individual selects a total order compatible with the partial order and evaluates it in the standard optimality-theoretic fashion.

Stress retraction, no contraction

how is it 3 2 1	*S/1	NSC	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) how is IT	1!		1	1	1		3
(b) how's IT	1!		1	1	1	1	2
(c) ☞ how IS it		1	1	1	1		3
(d) HOW is it		2!		1	1		3

Variable contraction

how is that 3 2 2	*S/1	NSC	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ how is THAT			1	1	1		3
(b) ☞ how's THAT			1	1	1	1	2
(c) how IS that		1	1	1	1		3
(d) HOW is that		2!		1	1		3

Variable contraction

she will be 2 3 2	*S/1	NSC	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ she will BE			1	1	1		3
(b) ☞ she'll BE			1	1	1	1	2
(c) she WILL be		1!		1	1		3
(d) SHE will be		2!	1	1	1		3

Content words (= Class 4) pose a problem

i have lee 2 4 4	*S/1	NSC	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ i have LEE					1		3
(b) ☞ i've LEE					1	1	2
(c) i HAVE lee		1			1		3
(d) I have lee		2	1	1	1		3

Solution: Indexed faithfulness (FAITH/n)

i have lee 2 4 4	FAITH/4	*S/1	NSC	*S/12	*S/123	*S/1234	FAITH	*WORD
(a) ☞ i have LEE						1		3
(b) i've LEE	1!					1	1	2
(c) i HAVE lee			1			1		3
(d) I have lee			2	1	1	1		3

A partial order for English phrasal stress

*S/1 >> NSC

*S/1 >> *WORD

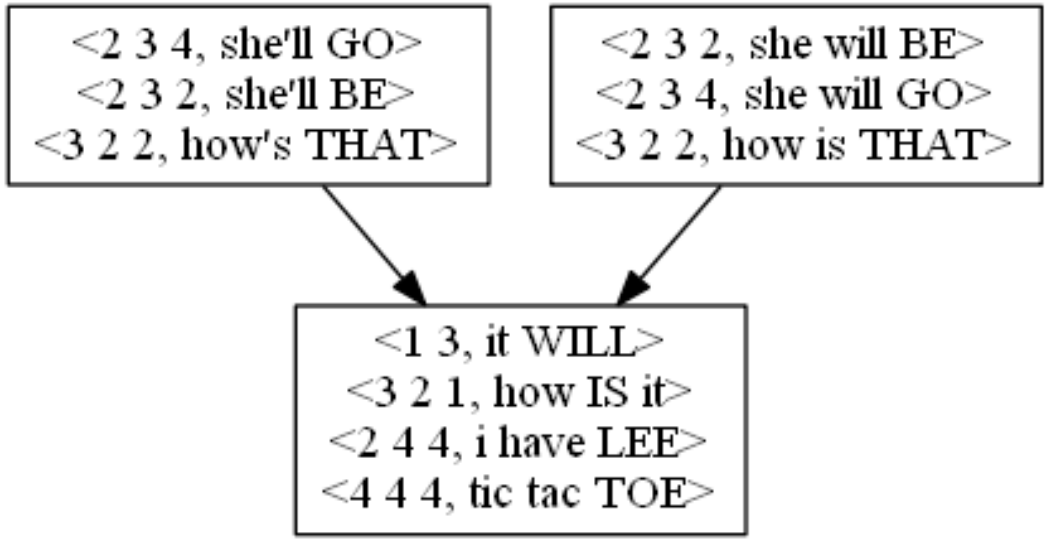
NSC >> *S/12

FAITH/4 >> *WORD

The predicted typology (phrasal stress, contraction)

		Output #1	Output #2	Contraction
(a)	/4 4 4/:	tic tac **TOE**	tic tac **TOE**	no
(b)	/1 3/:	it **WILL**	it **WILL**	no
(c)	/2 3 4/:	she will **GO**	she'll **GO**	variable
(d)	/3 2 1/:	how **IS** it	how **IS** it	no
(e)	/3 2 2/:	how is **THAT**	how's **THAT**	variable
(f)	/2 3 2/:	she will **BE**	she'll **BE**	variable
(g)	/2 4 4/:	i have **LEE**	i have **LEE**	no

ERC entailments (= T-order)



Empirical testing

The Buckeye Corpus of American English (Pitt et al. 2007)

- naturalistic speech, 40 speakers from Columbus, OH
- richly annotated, additional annotation by Sam Bowman
- focused on *will/shall*
- 769 relevant tokens: 533 contractions (*'ll*), 236 full forms (*will*, *shall*),
- 561 potentially variable tokens after exclusions

The right contexts of *will/shall* in Buckeye

(a) Monosyllabic function words (109): be, for

(b) Monosyllabic content words (379): all, ask, beat, bet, blow, break, buy, call, cause, change, chew, choose, claim, come, cost, count, deal, die, do, draft, draw, drive, ease, eat, end, feel, find, fit, flop, flunk, fool, get, give, go, have, hear, help, just, kind, know, lead, learn, leave, let, like, look, make, match, move, need, pay, pour, pull, put, raise, read, rent, save, say, see, send, set, share, shoot, show, sit, sleep, spend, start, stay, stick, still, stop, take, talk, tell, tend, then, they, think, try, turn, twist, use, vote, wait, wake, walk, watch, well, work, write

(c) Polysyllabic function words (0)

(d) Polysyllabic content words (75): actually, also, always, attack, basically, become, bury, continue, definitely, delete, depreciate, even, eventually, ever, expand, expect, explain, forget, happen, honor, ignore, listen, never, okay, only, order, organize, probably, protect, really, recognize, remember, repossess, retire, separate, suspend, tighten, usually, vacuum, wonder

The right contexts of *will/shall* in Buckeye

Our current analysis predicts no differences among these contexts:

- Only the weakest function words (Class 1) are predicted to allow phrasal stress retraction, blocking contraction.

The right contexts of *will/shall* in Buckeye

Our current analysis predicts no differences among these contexts:

- Only the weakest function words (Class 1) are predicted to allow phrasal stress retraction, blocking contraction.

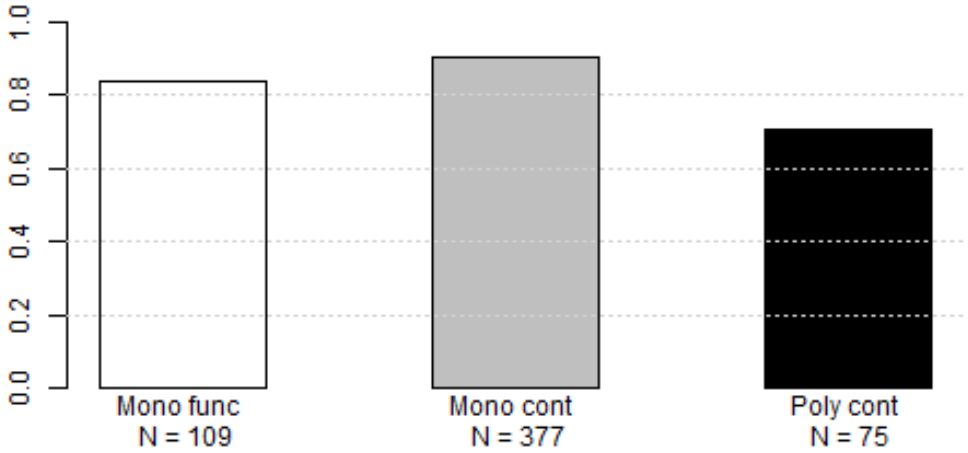
- All other function words (Class 2, Class 3) and all content words (Class 4) are predicted to attract phrasal stress off the auxiliary, allowing contraction.

The right contexts of *will/shall* in Buckeye

Our current analysis predicts no differences among these contexts:

- Only the weakest function words (Class 1) are predicted to allow phrasal stress retraction, blocking contraction.
- All other function words (Class 2, Class 3) and all content words (Class 4) are predicted to attract phrasal stress off the auxiliary, allowing contraction.
- The analysis predicts no difference between monosyllabic and polysyllabic right context words.

Contraction of *will/shall* in Buckeye by the right context



Why would a following polysyllable inhibit contraction?

- If the verb is monosyllabic, we get one binary phrase:

she will go → (*she'll GO*)

Why would a following polysyllable inhibit contraction?

- If the verb is monosyllabic, we get one binary phrase:

  she will go → (*she'll GO*)

- If the verb is longer, the result could be one ternary phrase or two binary phrases (Junko Itô, p.c.)

  she will explain → (*she'll exPLAIN*)
  she will explain → (*she WILL*) (*exPLAIN*)

Why would a following polysyllable inhibit contraction?

- If the verb is monosyllabic, we get one binary phrase:

  she will go → (*she'll GO*)

- If the verb is longer, the result could be one ternary phrase or two binary phrases (Junko Itô, p.c.)

  she will explain → (*she'll exPLAIN*)
  she will explain → (*she WILL*) (*exPLAIN*)

The latter puts *will* in a phrase-final position, blocking contraction.

Mixed-effects regression

Dependent variable: contraction vs. no contraction. Preceding consonant significantly disfavors and following monosyllable significantly favors contraction.

Random effects:

Groups	Name	Variance	Std.Dev.
speaker	(Intercept)	0.9858	0.9929
host.pron	(Intercept)	0.2020	0.4494

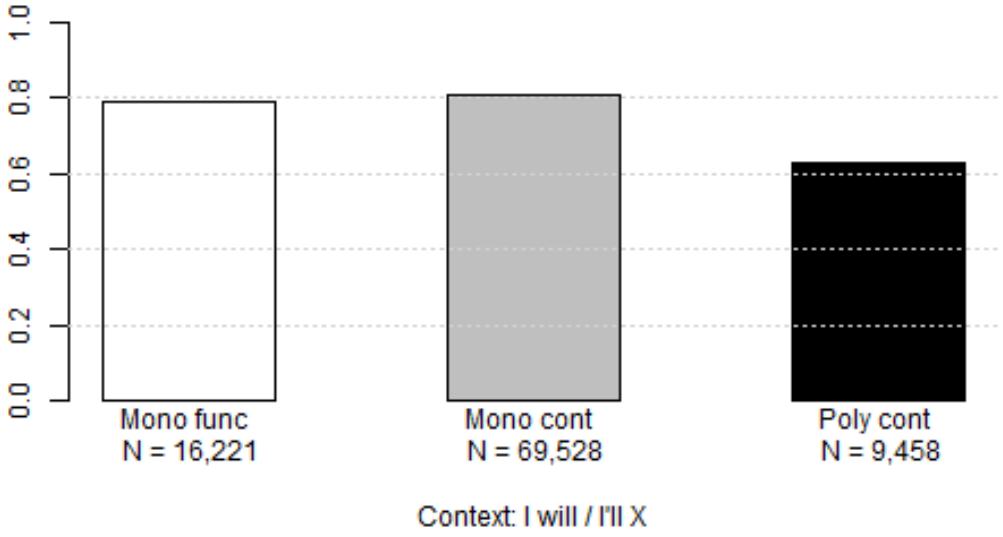
Number of obs: 561, groups: speaker, 39; hostword, 11

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.8604	0.6784	1.268	0.204717
prec.consTRUE	-1.7175	0.4876	-3.522	0.000428 ***
vowel.rate	0.1098	0.1025	1.072	0.283706
function.wordTRUE	-0.2811	0.3801	-0.740	0.459555
monosyllableTRUE	1.3099	0.3522	3.719	0.000200 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Contraction of *will/shall* in COCA (Davies 2008-) by the right context



OT analysis

Three constraints that strive to parse the input into binary phrases (see e.g., Itô and Mester 2003)

PARSE	'All syllables must belong to p-phrases'	
*MONO	'A p-phrase has at least two syllables'	(undominated)
*TERNARY	'A p-phrase has at most two syllables'	

OT analysis

Three constraints that strive to parse the input into binary phrases (see e.g., Itô and Mester 2003)

PARSE	'All syllables must belong to p-phrases'	
*MONO	'A p-phrase has at least two syllables'	(undominated)
*TERNARY	'A p-phrase has at most two syllables'	

Assumptions:

- Phrasal stress is by definition rightmost in a phrase.
- At most one syllable can be left unparsed.
- *MONO, FAITH/NSC, FAITH/4, *S/1 are undominated.

OT analysis

		NSC	*S/12	*S/123	*S/1234	*WD	FAITH	PARSE	*TERN
2 3 2	→ (she will BE)		1	1	1	3			1
	→ (she WILL) be	1		1	1	3		1	
	→ she (will BE)		1	1	1	3		1	
	→ (she'll BE)		1	1	1	2	1		
2 3 4	→ (she will GO)				1	3			1
	(she WILL) go	1		1	1	3		1	
	→ she (will GO)				1	3		1	
	→ (she'll GO)				1	2	1		
2 3 44	→ (she will exPLAIN)				1	3			1
	→ (she WILL) (exPLAIN)	1		1	2	3			
	she (will exPLAIN)				1	3		1	1
	→ (she'll exPLAIN)				1	2	1		1

Two predictions of the phrasing model

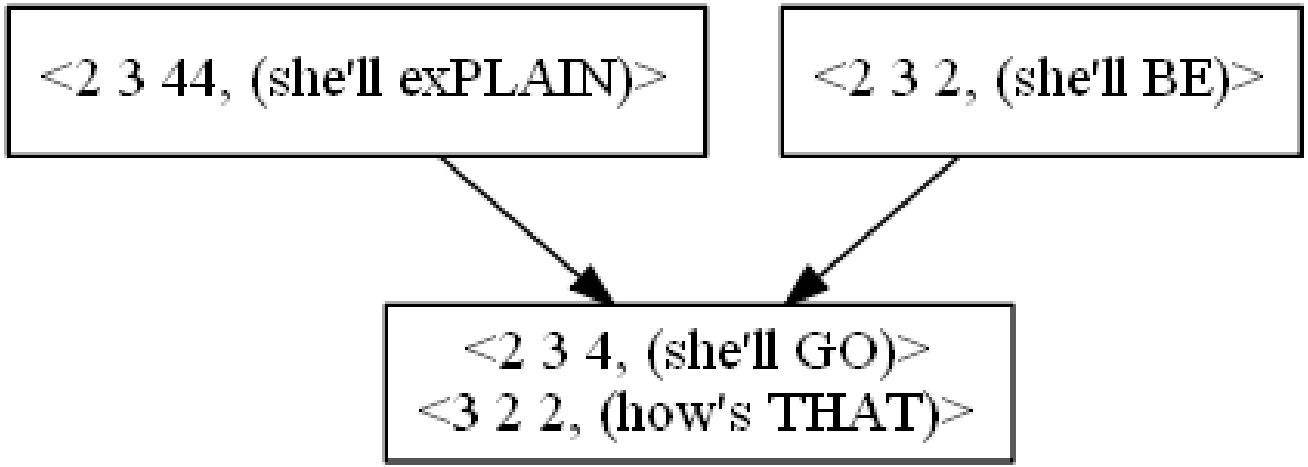
- Contraction is more natural before monosyllabic content words (*go*) than before polysyllabic content words (*explain*)

Two predictions of the phrasing model

- Contraction is more natural before monosyllabic content words (*go*) than before polysyllabic content words (*explain*)

- Contraction is more natural before content words (*go*) than before function words (*be*).

ERC entailments (= T-order), partial graph



An alternative explanation: UID

- **Uniform Information Density** (UID, Jaeger 2006, Levy and Jaeger 2006, Frank and Jaeger 2008:942): Speakers prefer choices that keep the amount of information uniform across the utterance.

A possible alternative explanation: UID

- **Uniform Information Density** (UID, Jaeger 2006, Levy and Jaeger 2006, Frank and Jaeger 2008:942): Speakers prefer choices that keep the amount of information uniform across the utterance.

- The information of a word is defined as the logarithm of the inverse of the probability of the word in its context.

A possible alternative explanation: UID

- **Uniform Information Density** (UID, Jaeger 2006, Levy and Jaeger 2006, Frank and Jaeger 2008:942): Speakers prefer choices that keep the amount of information uniform across the utterance.

- The information of a word is defined as the logarithm of the inverse of the probability of the word in its context.

- Polysyllabic words tend to be less frequent, hence high in information. Therefore speakers would prefer a full form of the auxiliary to avoid a spike in the rate of information transmission.

A possible alternative explanation: UID

Problem: *be*

- One would expect a high contraction rate before *be* because it is by far the most frequent next word (19% of all tokens) and hence low in information, but that is not what we find.

A possible alternative explanation: UID

Problem: *be*

- One would expect a high contraction rate before *be* because it is by far the most frequent next word (19% of all tokens) and hence low in information, but that is not what we find.

- Note that stress predicts the opposite: *be* should condition less contraction than content words. That is what we found.

More predictions of the phrasing theory of contraction

Three factors that determine phrasing (Gussenhoven 2004:159):

- SIZE: The length of prosodic constituents is subject to size constraints, e.g., binarity. Hence word length should play a role in contraction.
- FOCUS: A focused constituent tends to coincide with a prosodic constituent. Hence contraction should be blocked after focus.
- MORPHOSYNTAX: Prosodic constituents tend to coincide with morphosyntactic constituents. An auxiliary before a syntactic boundary should resist contraction.

Syntactic boundary effects

If phrasal stress is cyclic, a major syntactic boundary (more brackets) should block contraction more than a minor syntactic boundary (fewer brackets). Consider different adverbials:

They're tall, but I'm not. (Bender and Sag 2001:25)

(*i'm NOT*)

??Brad's very competitive, and I'm, too.

(*i AM*) (*TOO*) (Philip Spaelti, p.c.)

Syntactic boundary effects

Contraction frequencies from COCA: *just* vs. *then*

As for me, I'll just wait until spring.	(94.1%)
Well, then, I will just have to wait.	(5.9%)
If I’m in Maine, I’ll then do something with my family.	(9.4%)
Once all those things are in place, I will then do a line edit.	(90.6%)

Syntactic boundary effects

Contraction is blocked when the immediately following element has been deleted or displaced (e.g., Zwicky 1970, Baker 1971, Bresnan 1978, Kaisse 1983, Inkelas and Zec 1993, Wilder 1997):

Brad’s very competitive, and I am _ too.

Mary is a better lawyer than Sue is _ a doctor.

Tom is planting millet, and Lisa is _ peanuts.

I don't know where the party is _ tonight.

A major syntactic boundary between the auxiliary and the gap results in a phonological phrase boundary which blocks contraction).

Syntactic boundary effects

Contraction is

- disfavored before an NP
- favored before a verb, especially V-*ing* and *gonna*

  Labov 1969:731-732, McElhinny 1993, Sharma and Rickford 2009, MacKenzie 2012:166-171, Spencer 2014

In COCA, the average contraction rate of *will/shall* is

- 69.9% before *be* + NP (identified by *I'll be a/an/the*)
- 75.1% before *be* + a progressive verb (identified by *I'll be V-ing*)

  (p = 0.003247, Fisher's exact test)

Syntactic boundary effects

I am the moderator.	less contraction
I will be the moderator.	
I am talking with two experts.	more contraction
I will be talking with two experts.	

Why?

Syntactic boundary effects

Suggestion: Different syntactic structures result in different phrasings. Abstracting away from binarity, the following phrasings are predicted.

(i am TALKING) (with two EXPERTS)

(i will be TALKING) (with two EXPERTS)

(i AM) (the MODERATOR)

(i will BE) (the MODERATOR) ~ (i WILL be) (the MODERATOR)

Left context effects

(1a) I've gone there too often.

(1b) You'll like it in Manitoba.

(1c) You've painted your house.

(2a) *You and I've gone there too often.

(2b) *Grace and you'll like it in Manitoba.

(2c) *All the residents but you've painted their houses.

Left context effects

The longer the phrase, the less contraction.

(a) *The fact that it was she'll be a blow to the party.
(b) *The guy next to you'll speak first.
(c) *Anyone saying it was I'll be in big trouble
(d) *The two men who said it was they're arriving on the midnight plane.
(e) *A man as tall as he'll probably be shipped to Frederick the Great.
(f) *To see you'll be nice.
(g) *Everyone who hears you'll be impressed.

(examples from Zwicky 1970)

Copula contraction

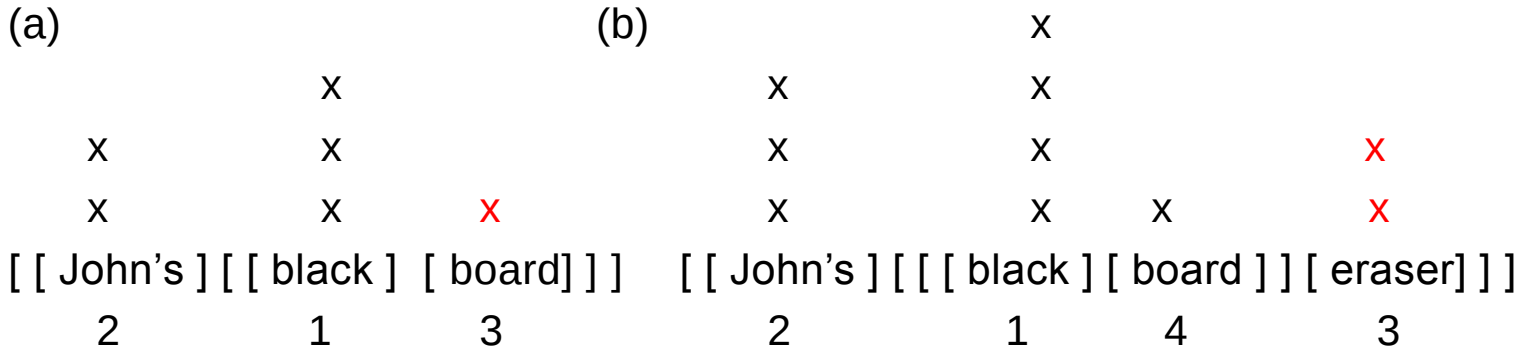
The subject length effect (MacKenzie 2012, Ch. 5):

(a) As subjects increase in length, contracted forms taper off.

(b) There are no contracted forms after subjects of more than eight words.

Possible explanation for the length effect

The more stress on the host word, the less eligible it is as a host.



Prediction

More contraction after a compound subject than a phrasal subject:

(a)	John’s BLACKboard is gone!	more contraction
(b)	John’s black BOARD is gone!	less contraction

Theoretical puzzles

- Spencer (2014) discovered that the phonetic duration of uncontracted copulas (e.g., *she is a student*) reflects the same contextual generalizations as the choice between uncontracted and contracted copulas (e.g., *she is ~ she's a student*) (cf. Halle's argument against the autonomous phoneme).

- Auxiliary contraction (i.e., allomorph selection) is sensitive to the phonological shape of the following word and the locus of phrasal stress. What does that tell us about locality?

Tentative conclusions

English Auxiliary Contraction depends on

- Word stress (four degrees)
- Phrasal stress
- Prosodic phrasing (binarity)

Much work remains to be done.

Thank you!