1. Preliminaries

(1) Auxiliary contraction is sometimes optional, sometimes blocked:
(a) you pay me i’ll do this thing  (b) *I think, therefore I’m
You’ll like it in Manitoba    *Grace and you’ll like it in Manitoba

(2) Syllabic, reduction  Non-syllabic, contraction
will [wɪɫ, wəl, œl]    ‘ll [l]
have [hæv, həv, əv]    ’v [v]

(3) Auxiliary contraction is POSTLEXICAL, i.e., sensitive to syntax and phrasal phonology. It is different from not-contraction (is not ~ isn’t) and to-contraction (want to ~ wanna) which are LEXICAL (Zwicky and Pullum 1983).


(5) Proposal: Contraction targets two metrically weak morphemes:
(a) Variation in contraction reflects variation in lexical and phrasal stress.
(b) Contraction is blocked if the host or the auxiliary carries phrasal stress.

(6) Both lexical and phrasal stress play a role:
(a) Auxiliaries contract (I’ve got a car), main verbs do not (*I’ve a car)
    because the latter are lexically stressed.
(b) Contraction is blocked under sentence stress (e.g., Yes, it is! vs. *Yes, it’s!)

(7) Outstanding descriptive puzzles:
(a) Lexical stress: Stress in monosyllabic function words is variable (see e.g., Selkirk 2003) and as a group they behave in a heterogeneous manner. For example, are shall, will, who, have, is, it lexically stressed or not?
(b) Phrasal stress: The Nuclear Stress Rule (Chomsky and Halle 1968,
    Liberman and Prince 1977) is a good first approximation of default phrase stress, but in reality there is plenty of variation.

(8) Plan of work:
(a) Adopt a preliminary analysis of lexical and phrasal stress.
(b) Adopt a hypothesis that connects stress to auxiliary reduction (= (5))
(c) Test the predictions against a corpus of naturalistic examples.
2. Theory

2.1 English stress

(9) Chomsky and Halle 1968, Liberman and Prince 1997:
   (a) The Nuclear Stress Rule (NSR): In a phrase (NP, VP, AP, S), assign stress to the rightmost word bearing lexical stress (= [1 stress]).
   (b) 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.

(10) The rules apply cyclically, starting from the innermost brackets, assigning stress to the relevant word and reducing stress elsewhere by one (stress subordination).

(11) 3 2 5 4 1

2.2 Optimality Theory

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

<table>
<thead>
<tr>
<th>tic tac toe</th>
<th>NSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) ( \rightarrow ) tic tac TOE</td>
<td>1</td>
</tr>
<tr>
<td>(b) tic TAC toe</td>
<td>2</td>
</tr>
<tr>
<td>(c) TIC tac toe</td>
<td></td>
</tr>
</tbody>
</table>

(13) Lexical stress: Morphemes enter into a STRESSABILITY HIERARCHY that plays a role in phrasal stress assignment and hence auxiliary contraction.

(14) Following earlier work (e.g., Lea 1979, O'Shaughnessy and Allen 1983, cited in Baart 1987; Ladd 1980, Altenberg 1987, Hirschberg 1993), we propose a tentative stressability hierarchy (Class 1 = weakest, Class 4 = strongest). We call the lexical stressability of a word its STRENGTH. For empirical diagnostics, see Shih 2013.

<table>
<thead>
<tr>
<th>CLASS</th>
<th>EXAMPLES</th>
<th>MORPHEME CLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( it, I, \text{them} )</td>
<td>personal pronouns</td>
</tr>
<tr>
<td>2</td>
<td>( \text{that}, \text{is}, \text{am, have}^{\text{AUX}} )</td>
<td>demonstratives, finite auxiliaries</td>
</tr>
<tr>
<td>3</td>
<td>( \text{could, will, how} )</td>
<td>modals, WH-words</td>
</tr>
<tr>
<td>4</td>
<td>( \text{gone, John, have}^{\text{LEX}} )</td>
<td>open class words</td>
</tr>
</tbody>
</table>
Lexical and phrasal stresses interact: the morpheme’s lexical stress class affects its ability to carry phrasal stress.

Markedness: Assign a violation mark for every morpheme of a given class that carries phrasal stress (for the stringency format, see e.g., Kiparsky 1994, Prince 1997, de Lacy 2004)

<table>
<thead>
<tr>
<th>Markedness</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*STRESS/1</td>
<td>No phrasal stress on Class 1.</td>
</tr>
<tr>
<td>*STRESS/12</td>
<td>No phrasal stress on Classes 1 or 2.</td>
</tr>
<tr>
<td>*STRESS/123</td>
<td>No phrasal stress on Classes 1 or 2 or 3.</td>
</tr>
<tr>
<td>*STRESS/1234</td>
<td>No phrasal stress on Classes 1 or 2 or 3 or 4.</td>
</tr>
</tbody>
</table>

Economy: Assign a violation mark for each word (*WORD).

Faithfulness: Assign a violation mark for every contracted morpheme (FAITH).

Simple as it is, this theory already makes a number of unobvious predictions.

2.3 Predictions

General predictions: Given an auxiliary,

(a) the stronger the next word, the more it attracts stress off the auxiliary, favoring contraction.
(b) the weaker the next word, the more stress tends to land on the auxiliary, disfavoring contraction.
(c) if there’s no following word at all (= phrase break) stress must fall on the auxiliary, blocking contraction categorically.

Example: Phrasal stress falls on the rightmost word except if there’s a stronger word in the left context in which case it may be minimally retracted.

Input = string of lexical stresses, e.g., how = 3, is = 2, it = 1
(b) Output = string of lexical stresses with nuclear stress added (in capitals)

In a dialect where *S/1 >> NSC >> the rest, if the phrase-final word is weaker than the auxiliary, (i) contraction is blocked; (ii) nuclear stress is minimally retracted.

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>*S/1</th>
<th>NSC</th>
<th>*S/12</th>
<th>*S/123</th>
<th>FAITH</th>
<th>*WORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>how is it = 3 2 1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(a) how is IT</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>(b) how’s IT</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) how IS it</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>(d) HOW is it</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

In (a) and (b) of the example, the auxiliary attracts nuclear stress if the following word is stronger;

How’s the weather in Boston? (Zwicky 1970)

*How’s it in Boston? (Zwicky 1970)

(c) Who’s it? (in a game, it = Class 4) (Labov 1969:722)
(25) In the same dialect, if the phrase-final word is of the same strength as the auxiliary, (i) contraction is optional; (ii) nuclear stress is final.

<table>
<thead>
<tr>
<th>how is that = 3 2 2</th>
<th>*S/1</th>
<th>NSC</th>
<th>*S/12</th>
<th>*S/123</th>
<th>FAITH</th>
<th>*WORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) how is THAT</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>(b) how’s THAT</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>(c) how IS that</td>
<td>1!</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>(d) HOW is that</td>
<td>2!</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

(26) Variation (e.g., Kiparsky 1993, Anttila and Cho 1998, Zamma 2012, Djalali 2013):  
(a) Grammars are partially ordered sets of constraints, in this case  
   *S/1 >> NSC >> {*S/12, *S/123, FAITH, *WORD}  
(b) At the moment of performance, an individual selects a total order compatible with the partial order (24 total orders in all).

(27) In the same dialect, there’s no contraction or stress retraction if the auxiliary is phrase-final, irrespective of the host’s strength:

<table>
<thead>
<tr>
<th>it will = 1 3</th>
<th>*S/1</th>
<th>NSC</th>
<th>*S/12</th>
<th>*S/123</th>
<th>*S/1234</th>
<th>FAITH</th>
<th>*WORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) it WILL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(b) IT will</td>
<td>1!</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kit will = 4 3</th>
<th>*S/1</th>
<th>NSC</th>
<th>*S/12</th>
<th>*S/123</th>
<th>*S/1234</th>
<th>FAITH</th>
<th>*WORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Kit WILL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(b) KIT will</td>
<td>1!</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

(28) Note that IT will has more violations than KIT will. This will matter in variation.

(29) In the same dialect, contraction is optional if the auxiliary is not phrase-final and the final word is stronger than the auxiliary:

<table>
<thead>
<tr>
<th>he will go = 1 3 4</th>
<th>*S/1</th>
<th>NSC</th>
<th>*S/12</th>
<th>*S/123</th>
<th>*S/1234</th>
<th>FAITH</th>
<th>*WORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) he will GO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(b) he’ll GO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(c) he WILL go</td>
<td>1!</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(d) HE will go</td>
<td>1!</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

(30) More evidence: not lures phrasal stress off the auxiliary, making contraction possible (examples from Bender and Sag 2001:25):  
(a) *They’re not tall, but I’m.  
   They’re tall, but I’m not.  
(b) *Kim isn’t coming, but the student I met’s.  
   Kim’s coming, but the student I met’s not.
2.4 Evidence for split faithfulness

(31) Lexical verbs, e.g., have\textsubscript{LEX} (Class 4) do not contract. If the object is weak (Class 1), we happen to get the facts right:

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{i have them} & 1 & 4 & 1 & *S/1 & NSC & *S/12 & *S/123 & *S/1234 & \text{FAITH} & *\text{WORD} \\
\hline
(a) i have THEM & 1! & 1 & 1 & 1 & 1 & 1 & 1 & 3 \\
(b) i've THEM & 1! & 1 & 1 & 1 & 1 & 1 & 2 & 3 \\
\rightarrow (c) i HAVE them & 1 & 1 & 1 & 1 & 1 & 3 \\
(d) I have them & 1! & 2 & 1 & 1 & 1 & 3 \\
\hline
\end{array}
\]

(32) Things go wrong if the object is strong (Class 4):

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{i have lee} & 1 & 4 & 4 & *S/1 & NSC & *S/12 & *S/123 & *S/1234 & \text{FAITH} & *\text{WORD} \\
\hline
\rightarrow (a) i have LEE & 1 & 1 & 1 & 1 & 3 \\
\rightarrow (b) i've LEE & 1 & 1 & 1 & 2 \\
(c) i HAVE lee & 1! & 1 & 1 & 3 \\
(d) I have lee & 1! & 2 & 1 & 3 \\
\hline
\end{array}
\]

(33) Solution: Special faithfulness

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{i have lee} & 1 & 4 & 4 & *S/1 & NSC & *S/12 & *S/123 & *S/1234 & \text{FAITH} & *\text{WORD} \\
\hline
\rightarrow (a) i have LEE & \text{FAITH/4} & 1 & 1 & 1 & 1 & 3 \\
(b) i've LEE & 1! & 1 & 1 & 2 \\
(c) i HAVE lee & 1! & 1 & 3 \\
(d) I have lee & 1! & 2 & 3 \\
\hline
\end{array}
\]

(34) This suggests that there exists a hierarchy of faithfulness constraints that mirrors the hierarchy of stress constraints.

(35) FAITH/4 No contraction in Class 4.
FAITH/43 No contraction in Classes 3 or 4.
FAITH/432 No contraction in Classes 2 or 3 or 4.
FAITH/4321 No contraction in Classes 1 or 2 or 3 or 4.

(36) In other words, the stressability of a morpheme is inversely related to its ability to undergo contraction. This seems right: it is well known that is (Class 2) contracts in a much larger set of contexts than does will (Class 3) (Spencer 2013).
2.5 Quantitative predictions

(37) What does this theory predict about quantitative patterns in variable cases?

(38) Comparing three different inputs:

<table>
<thead>
<tr>
<th>Lexical Stress</th>
<th>*S/1</th>
<th>*S/12</th>
<th>*S/123</th>
<th>*S/1234</th>
<th>WORD*</th>
<th>NSC</th>
<th>FAITH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>SHE will be</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she WILL be</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she will BE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she'll BE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 3 4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>SHE will go</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she WILL go</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she will GO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she'll GO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SHE will</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she WILL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(39) Given these three inputs, our constraints derive a factorial typology of four possible dialects, computed by OTSoft (Hayes et al. 2003) below. No rankings are assumed. Contraction is shaded.

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Grammar #1</th>
<th>Grammar #2</th>
<th>Grammar #3</th>
<th>Grammar #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3 2</td>
<td>she WILL be</td>
<td>she WILL be</td>
<td>she will BE</td>
<td>she'll BE</td>
</tr>
<tr>
<td>1 3 4</td>
<td>she will GO</td>
<td>she'll GO</td>
<td>she will GO</td>
<td>she'll GO</td>
</tr>
<tr>
<td>1 3</td>
<td>she WILL</td>
<td>she WILL</td>
<td>she WILL</td>
<td>she WILL</td>
</tr>
</tbody>
</table>

(40) No dialect contracts phrase-finally.

(41) Implicational universals: In all dialects in the factorial typology, contraction in she’ll be implies contraction in she’ll go, but not vice versa.

(42) Variable dialects are mixtures of the four invariant dialects. Examples:
(a) Grammar {#1, #2} → she will GO ~ she’ll GO
(b) Grammar {#1, #2, #3, #4} → she WILL be ~ she will BE ~ she’ll BE; she will GO ~ she’ll GO

(43) Quantitative prediction: In dialect {#1, #2, #3, #4} contraction should be more frequent in she’ll go (= 2/4) than in she’ll be (= 1/4) assuming that an individual selects a total ranking randomly from the partial order at the moment of performance (Kiparsky 1993, Anttila 1997, Anttila and Cho 1998, Riggle 2010)

(44) Theoretical upshot: Quantitative patterns arise from IMPLICATIONAL UNIVERSALS hidden in the grammar.
(45) Implicational universals visualized (T-order, Anttila and Andrus 2006)

(46) Interim summary = (20): Given an auxiliary,
   (a) the stronger the next word, the more it attracts stress off the auxiliary, favoring contraction.
   (b) the weaker the next word, the more stress tends to land on the auxiliary, disfavoring contraction.
   (c) if there’s no following word at all (= phrase break) stress must fall on the auxiliary, blocking contraction categorically.

3. Empirical tests

(47) Data: will / shall / ’ll in the Buckeye Corpus of American English (Pitt et al. 2007). 40 speakers, 769 tokens, 533 contracted, 236 uncontracted. Thanks to Sam Bowman for the Python script for extracting and annotating the data.

(48) We excluded cases where blocking is near-categorical (e.g., left context is a conjunction, silence, noise or a non-pronoun; right context is the end of a phrase; disfluencies). This leaves us with 563 tokens.

3.1 Right context

(49) Prediction: The stronger the next word, the more it should attract stress off the auxiliary, hence the more auxiliary contraction we should observe.

(50) Unfortunately, not enough data to test all four stress classes. Instead, we tested the following related hypotheses: (i) Content words should be stronger than function words; (ii) Polysyllabic words should be stronger than monosyllabic words.
(51) Monosyllabic function words (109): *be, for*

(52) 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, uh, use, vote, wait, wake, walk, watch, well, work, write*

(53) Polysyllabic function words (0)

(54) 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*

(55) Data from the Buckeye Corpus (Pitt et al. 2007). Raw frequencies (N = 563):

![Contraction of "will/shall" in the Buckeye Corpus](image)

(56) Monosyllabic function words (*be*, class 3) vs. monosyllabic content words (class 4) trend in the expected direction, but not significant (Fisher, p = 0.08643).

(57) **Surprise**: Significantly less contraction before polysyllables than before monosyllables (p = 4.329e-05).

(58) The effect survives under various controls.
A mixed-effects logistic regression model implemented in package lme4 of R (R Development Core Team 2007). Fixed effects: preceding consonant, speech rate (vowels per second), following function word, following monosyllable, following verb. Random effects: speaker (39), host (11). Model summary:

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.f0</td>
<td>(Intercept)</td>
<td>0.92680</td>
<td>0.96271</td>
</tr>
<tr>
<td>hostdict</td>
<td>(Intercept)</td>
<td>0.17955</td>
<td>0.42374</td>
</tr>
</tbody>
</table>

Number of obs: 563, groups: f.f0, 39; hostdict, 11

Fixed effects:

| Estimate | Std. Error | z value | Pr(>|z|) |
|----------|------------|---------|----------|
| (Intercept) | 1.1744     | 0.6865  | 1.711    | 0.08712 . |
| precCTRUE   | -1.7446    | 0.4621  | -3.775   | 0.00016 *** |
| vrate       | 0.1061     | 0.1021  | 1.040    | 0.29850 |
| function.wordTRUE | -0.2225 | 0.3758  | -0.592   | 0.55389 |
| monosyllableTRUE | 1.6518    | 0.4062  | 4.067    | 4.76e-05 *** |
| next.word.verbTRUE | -0.6918 | 0.4463  | -1.550   | 0.12115 |

---

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

What might explain this?

Uniform Information Density (UID, Frank and Jaeger 2008:942): Full form before a low-frequency word helps lexical access. Problem: be is by far the most frequent next word (19% of all tokens), but does not result in the highest contraction rate.

Phrasing: Long verbs are phrased separately from the host + auxiliary, e.g., you will go → (you’ll GO) vs. you will remember → (you WILL) (reMEMber). Analysis: Replace NSC with constraints that strive for BINARITY (Itô and Mester 2003) Assumptions: (i) nuclear stress rightmost; (ii) at most one unparsed syllable

<table>
<thead>
<tr>
<th>PARSE</th>
<th>‘All syllables must belong to p-phrases’</th>
</tr>
</thead>
<tbody>
<tr>
<td>*MONO</td>
<td>‘A p-phrase has at least two syllables’</td>
</tr>
<tr>
<td>*TERNARY</td>
<td>‘A p-phrase has at most two syllables’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>*1</th>
<th>*2</th>
<th>*3</th>
<th>*4</th>
<th>*1</th>
<th>*2</th>
<th>*3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3 4</td>
<td>(she will GO)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(she WILL) go</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she (will GO)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(she'll GO)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>*1</th>
<th>*2</th>
<th>*3</th>
<th>*4</th>
<th>*5</th>
<th>*6</th>
<th>*7</th>
<th>*8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3 4</td>
<td>(she will EXPLAIN)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(she WILL) (EXPLAIN)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>she (will EXPLAIN)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(she'll EXPLAIN)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This grammar predicts the observed asymmetry (partial graph):

This grammar predicts the observed asymmetry (partial graph):

![Diagram showing contraction of "I will X" in COCA Corpus]

Data from the COCA Corpus (Davies 2008-). Raw frequencies of “I will X”:

What does the analysis predict about contraction in the following sentences?
(a) I will be the moderator
(b) I will be talking with two experts

Suppose NPs and PPs are phonological phrases (Truckenbrodt 2007) and VPs otherwise uniformly right-branching. The following phrasings will be favored:
(a) I (will be) (the moderator)
(b) I (will be talking) (with two experts)

Predicted stresses under normal accentuation, abstracting away from binarity:
(a) I (will BE) (the MODERATOR) ~ I (WILL be) (the MODERATOR)
(b) I (will be TALKING) (with two EXPERTS)
(68)  
(a) In (a), will gets phrasal stress variably because its only competitor is weaker (be, class 2) and there is no stronger word within the same phrase.  
(b) In (b), will is not eligible for phrasal stress because there is a stronger word within the same phrase (talking, class 4).  

(69)  
Prediction: Less contraction before NP (= (a)) than before V-ing (= (b)). In COCA, we have 70% (I’ll be a/an/the) and 76% (I’ll be V-ing) (p = 0.0003). See also McElhinny 1993, Sharma and Rickford 2009, Spencer 2013.  

3.2 Left context  

(70)  
Sadler (1997): Nonsyllabic contracted auxiliaries are inflectional affixes, not clitics.  

(71)  
Argument from coordination: Nonsyllabic contracted auxiliaries fail to take scope over a coordinate subject, e.g., *John and you’ll go home, *You and I’ve already tried to help Kim. This is expected of an affix, which is word-internal.  

(72)  
Bender and Sag (2001, 26): “such forms […] would involve the coordination of a noun (you) and a verb (I’ve), which is generally disallowed in coordination.”  

(73)  
This analysis seems incompatible with the stress hypothesis: phrasal stress would have to influence the choice of allomorphs within a word (cf. MacKenzie 2012:56).  

(74)  
But blocking extends beyond coordination (Zwicky 1970):  
(a)  
*You and I’ve gone there too often.  
*Grace and you’ll like it in Manitoba.  
*All the residents but you’ve painted their houses.  
(b)  
*The fact that it was she’ll be a blow to the party.  
*The guy next to you’ll speak first.  
*Anyone saying it was I’ll be in big trouble  
*The two men who said it was they’re arriving on the midnight plane.  
*A man as tall as he’ll probably be shipped to Frederick the Great.  
(c)  
*To see you’ll be nice.  
*Everyone who hears you’ll be impressed.  

(75)  
The blockers are morphosyntactically heterogeneous, but phonologically a natural class: the final word receives phrasal stress except perhaps not in (c).  

(76)  
Assume phrasal stress is rightmost and at most one syllable can be extrametrical. The blocking of contraction falls out from binarity:  
(a)  
(you’ll LIKE) it  
(b)  
(Grace and YOU)(will LIKE) it  

(77)  
The subject length effect (MacKenzie 2011, 2012):  
(a)  
As subjects increase in length, contracted forms taper off.  
(b)  
No contracted forms after subjects of more than eight words.
MacKenzie: Such length sensitivity is uncharacteristic of grammatical alternations and suggests that extragrammatical, memory-based constraints are at work.

This may be a stress effect in disguise: the taller the stress column, the less eligible the host (cf. Katz and Selkirk 2011):

(a)             (b)       x
       x       x       x
       x       x       x       x
  [ [ John’s ] [ [ black ] [ board ] ] ]  [ [ John’s ] [ [ [ black ] [ board ] ] [ eraser ] ] ]

Contraction should be easier after a compound than after a phrase of the same length, e.g., *John’s BLACKboard’s gone vs. John’s black BOARD’s gone.

4. Constituent deletion and displacement

Contraction is blocked when the immediately following element has been deleted or displaced (e.g., Zwicky 1970, Baker 1971, Bresnan 1978, Kaisse 1983) as if the auxiliary were sentence-final.

(a) Mary is a better lawyer than Sue is a doctor.
(b) No WAY is he gonna do that!

Tentative proposal: In (a), ellipsis entails a phonological phrase boundary after the auxiliary (Inkelas and Zec 1993, Wilder 1997), resulting in phrasal stress on is. In (b), contraction may be blocked by focus stress on the host.

5. Frequency

The joint frequency of host and auxiliary may well be a factor (Krug 1998, Spencer 2013), but not the only one: I’ll go, *Grace and I’ll go, *Yes, I’ll.

6. Conclusion

Many authors acknowledge the role of prosody in auxiliary contraction, but provide no analysis, focusing on other factors such as syntax instead.

If we do not understand the prosody, we cannot control for it. Once we do understand it, maybe nothing more is needed (Selkirk 1984:406).

If this theory is on the right track, auxiliary reduction provides a convenient diagnostic for stress and phrasing, applicable to both spoken and written data.
Acknowledgements

Thanks to audiences at the *Workshop on Locality and Directionality at the Morphosyntax-Phonology Interface* (Stanford University, October 12-14, 2012), the Michigan State University (May 2, 2013), and the *International Workshop on the Grammatical Word* (UC Davis, May 10, 2013), especially Raúl Aranovich, Robert Bayley, Sam Bowman, Joan Bresnan, Cleo Condoravdi, Kayla Cornale, Karthik Durvasula, David Embick, Heidi Harley, Grover Hudson, Sharon Inkelas, Junko Itô, Paul Kiparsky, Martha McGinnis, Armin Mester, Alan Munn, Owen Rambow, Stephanie Shih, Cristina Schmitt, Jessica Spencer, Peter Svenonius, and Draga Zec.

References


Kiparsky, Paul. 1994. ‘Remarks on markedness’, Handout from TREND (Trilateral Phonology Weekend), Santa Cruz.


Spencer, Jessica. 2013. ‘Multi-level effects on morphosyntactic variation: toward a model of copula contraction’, Dissertation Proposal, Department of Linguistics, Stanford University.


anttila@stanford.edu