Expanding StressTyp2:
Integrating Pitch Accent and Tone

Adam Jardine and Amanda Payne
Department of Linguistics and Cognitive Science, University of Delaware

INTRODUCTION

- **StressTyp2 (ST2)** (st2.ullet.net) is a new database of stress and accent patterns in over 700 languages based on two existing databases: Jeffrey Heinz's **Stress Pattern Database (SPD)** and Rob Goedemans and Harry van der Hulst's **StressTyp Database (ST1)**. The creation of ST2 is supported by grant no. 1123692 from the National Science Foundation
- ST2 catalogues patterns of secondary and primary stress, but includes little information on pitch accent or tonal systems
- We propose a way of expanding the structure of ST2 to include pitch accent and (limited) tone systems

STRESSTYP2

- MySQL database containing information on:
  - Languages (dialects, geographical information, endangerment)
  - Patterns (primary, secondary, compound,...)
  - Theories (parametric, computational,...)
  - Sources
- 760 (and counting!) languages and 918 unique patterns
- Data collected from a variety of descriptive and theoretical sources

ACCENT & TONE

- How to expand to pitch accent and tone systems?
- Two key differences from predictable stress patterns:
  - Pitch accent and tone can be lexically specified
  - Mapping between UR and SR is often important

DESCRIBING PATTERNS IN STRESSTYP2

- ST2 has three, atheoretic ways to describe surface stress patterns:

- **Stress Priority Codes** (**SPCs**; Bailey 1995, Heinz 2009)
  - Numbers mark position of prioritized syllables
  - L or R mark direction

- **StressTyp Codes** (**STCs**; van der Hulst and Goedemanns)
  - Alphabetic abbreviations for stress location
  - Weight classes shown by "/"

- **Metrical parameters**
  - Parameters for both primary & secondary stress

- **Finite State Automata** (**FSAs**)
  - Computational formalism describing (potentially infinite) sets of strings
  - Used to study complexity properties of patterns (as in Heinz 2009)

Figure 1: Hopi on the StressTyp2 website

Hopi: "In words of all sizes, primary stress falls on the initial syllable if it is heavy, else on the peninitial syllable if it is heavy, else on the peninitial syllable."

Hopi Parameters:

Domain: Left
Weight sensitive: Yes
Stress if both heavy: Left
Stress if both light: iambic
Heavy for stress: Long vowels; closed syllables

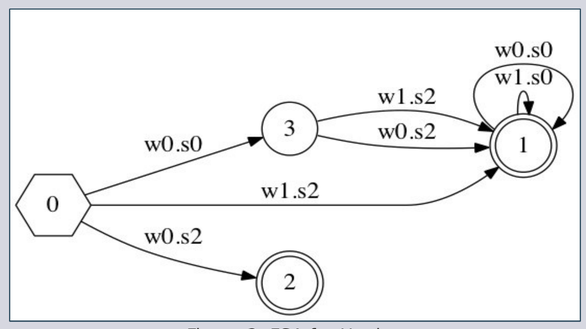

Figure 2: FSA for Hopi

KAGOSHIMA JAPANESE

- A straightforward example is Kagoshima Japanese
- A 2-pattern system, the accent falls on either the ultimate or penultimate syllable (Ishihara 2012)

Type A (penutimate)		Type B (ultimate)	
HL	hana 'nose'	LH	hana 'flower'
LHL	sakura 'cherry blossom'	LLH	usagi 'rabbit'
LLHL	kagaribi 'watch fire'	LLLH	kakimono 'document'

Kagoshima Japanese in StressTyp2:

SPC: 2R or 1R **FSA:**

STC: U or I

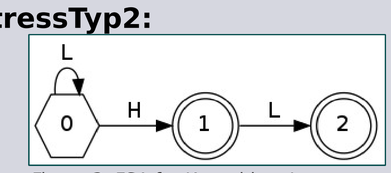

Figure 3: FSA for Kagoshima Japanese

A BANTU TONE SYSTEM

- Can we expand this to (limited) tone systems?
- Digo (Kisseberth 1984) is a Bantu tone language with only a H/Ø distinction
- If one H, it appears on the final mora of the phrase

kuhenza	'look for'	LLL	kuhenza muganga	'~ a doctor'	LLL LLL
kuiha	'call'	LLH	kuiha muganga	'~ a doctor'	LLL LLH

anahenza 'he looks for' LLLH
anahenza muganga 'he looks for a doctor' LLLL LLH

- More than one tone is possible; this leads to complex behavior!

Simple Digo in StressTyp2:

SPC: 1R **Parameters:**
STC: U
Culminative: No
Obligatory: No
Domain: Phrase

FST:

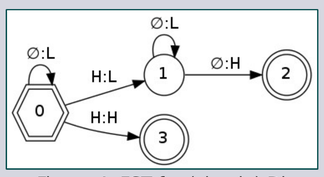

Figure 4: FST for (simple) Digo

TOKYO JAPANESE

- Tokyo Japanese (McCawley 1968, Pierrehumbert & Beckman 1988, and many others) is less straightforward
- Position of accent is lexically specified (Kubozono 2008 argues for 'default' position)
- Many words are unaccented
- 'Accent' is manifested as a sequence of H tone morae followed by a drop to L

Trisyllabic nouns with case particle (Kubozono 2012)

HLL-L	inoti-ga	'life-NOM'
LHL-L	kokoro-ga	'heart-NOM'
LHH-L	otoko-ga	'man-NOM'
LHH-H	sakana-ga	'fish-NOM'

Tokyo Japanese in StressTyp2:

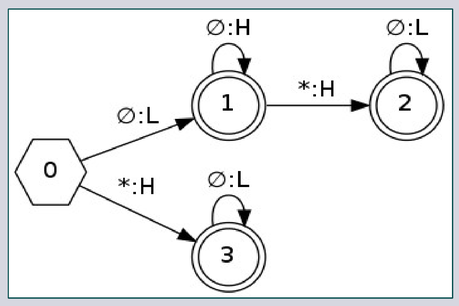

Figure 4: FST for Tokyo Japanese

- **Finite-state transducer** showing how accent behaves (Figure 4)

- **SPC and STC** describing the default pattern (Kubozono 2008):
  "Accent is assigned to the syllable containing the third mora from the right"
  SPC: 3R(μ) **STC:** A(μ)

- **Parameters** based on those discussed in Hyman (2006), Uwano (2012) and Kubozono (2012):
  Culminative: Yes
  Obligatory: No
  Unit for counting: Mora
  Unit for assignment: Syllable
  Number of patterns: n+1 (where n is number of syllables)
  Registers: None
  Kernel type: Lowering

ACKNOWLEDGEMENTS

We would like to thank ST2 PIs Jeffrey Heinz and Harry van der Hulst, as well as Dr. Haruo Kubozono for his assistance with an application for the NSF East Asia and Pacific Summer Institute, a grant that would help realize the ideas outlined in this poster. ST2 is made possible by NSF Grant #1123692.

CONCLUSIONS/FURTHER WORK

- Preliminary steps towards larger database of stress, pitch accent, and tone
- Targets exactly the kind of stress-like pitch accent and tone systems that are relevant to questions about typology of stress, pitch accent, and tone (Hyman 2009, van der Hulst 2011)
- How to deal with systems with interactions between more than one tone?
  - One FST per 'rule'; complex FST describing whole system
  - Parameters governing tone interactions (Meussen's Rule, boundedness, etc.)
- How to encode autosegmental representations?
- How to incorporate contour tone systems, like Mandarin Chinese?

REFERENCES

Bailey, T. M. 1995. Nonmetrical Constraints on Stress. Doctoral dissertation, University of Minnesota. Ann Arbor, MI: UMI
Goedemans, R., J. Heinz, and H. van der Hulst. 2013. StressTyp2. Forthcoming.
Gordon, M. 2002. A factorial typology of quantity insensitive stress, 2002, *Natural Language and Linguistic Theory* 20, 491-552
Heinz, J. 2009. On the role of locality in learning stress patterns. Phonology, 26(2):303-351, 2009
Hyman, L.M., 2006. Word-prosodic typology. Phonology 23, 225—257.
Hyman, L. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31: 213-238.
Ishihara, S. 2012. Osaka and Kagoshima Japanese citation tone acoustics. *Journal of the International Phonetic Association*. 42(1): 1-21.
Kisseberth, C. 1984. Digo tonology. In Clements & Goldsmith, Autosegmental Studies in Bantu Tone: 105–182. Foris Publications.
Kubozono, H. 2008. Japanese accent. In: Miyagawa, S., Saito, M. (Eds.), Handbook of Japanese Linguistics. Oxford University Press, Oxford, pp. 165—191.
Kubozono, H. 2012. Varieties of pitch accent systems in Japanese. Lingua, 122. 1395–1414.
McCawley, J.D., 1968. The Phonological Component of a Grammar of Japanese. The Hague & Paris, Mouton.
Pierrehumbert, J.B., Beckman, M.E., 1988. Japanese Tone Structure. MIT Press, Cambridge, MA.
Uwano, Z. 2012. Three types of accent kernels in Japanese. *Lingua*. 122. 1415–1440.
van der Hulst, H., R. Goedemans and E. Van Zanten. 2010. A survey of word accentual patterns in the languages of the world. Mouton de Gruyter: Berlin
van der Hulst, Harry (2011). Pitch accent systems. In Marc van Oostendorp et al. (ed.), *PitchAccent Systems*, Volume 3, 1003–27. Blackwell.