A dynamical systems approach to F0 hard-landing downtrends in Embosi

Yubin Zhang[1], Annie Rialland[2] & Louis Goldstein[1]

[1]University of Southern California [2]Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Sorbonne-Nouvelle, 4 rue des Irlandais, 75005 Paris, France
yubinzha@usc.edu, annie.rialland@sorbonne-nouvelle.fr, louisgol@usc.edu

Dynamical systems have increasingly come to be employed in the representation of speech, leveraging the lawful relation they provide between the discrete, context-free parametric specification of a system (for example a phonetic unit) and the continuous, context-dependent change in some measurable quantity [1]. Articulatory Phonology [2], [3] initially advanced this dynamical perspective by proposing that the primitive units of phonological representations and speech production are dynamically specified vocal tract constriction gestures. Later developments of this particular dynamical framework began to model prosody—both phasal boundaries and intonation—by incorporating prosodic planning dynamics [4]. The current work extends the dynamical systems approach to the investigation of utterance-level intonational events.

We focus on the utterance-level f0 downtrend phenomena, where the f0 of speech declines over an utterance, especially for declaratives. F0 downtrends have been theorized to arise from different intonational events, such as more global components like f0 downstep (downdrift) and declination, and more localized f0 final lowering [5]. Existing quantitative models of downstep, built under the assumption of separation of phonological units from quantitative f0 measurements, have included an exponential decay function to model the f0 patterns (e.g, Liberman & Pierrehumbert, 1984; Myers, 1996). On a dynamical view, these models are equivalent to a first-order system with point-attractor dynamics (whose analytic solution is an exponential decay). One problem for this model is that not all languages exhibit global intonational trends that can be readily characterized by point-attractor dynamics alone. Laniran & Clements (2003) use the terms soft landing versus hard landing to describe two types of f0 global trends with different kinematic profiles. In soft-landing like the attractor dynamics, the f0 approaches an asymptote smoothly with near zero velocity, whereas, in hard landing, f0 drops quickly and the negative velocity is relatively large utterance-finally. Laniran & Clements (2003) postulate a linear model to account for some of the hard-landing patterns found in their Yoruba data. An even more extreme pattern of hard landing seems to be present in Embosi, a Bantu language (C25) spoken in the Republic of Congo [8]. Initial observations suggest that the f0 contour in declarative sentences in Embosi exhibits initial rising and final hard landing. This suggests a distinct type of dynamical system, namely the free-fall dynamics that models the parabolic flight trajectory of gravitational fields. The current study analyzes f0 downtrends in a corpus of Embosi data and shows that it can be well modelled using primarily a free-fall-style dynamical system.

The utterances analyzed in the current study were taken from an Embosi corpus recorded by three native speakers of Embosi [8], [9]. The F0 contour of each utterance was extracted using the auto-correlation algorithm implemented in `Praat`. The mean f0 of each target moraic interval was calculated. The initial and final f0 events were analyzed separately using mixed-effects models. The statistical results reaffirm that the utterance-level f0 contour in Embosi declaratives resembles a parabolic curve, exhibiting initial rising and final 'hard-landing' patterns (See Fig. 1). We also found evidence for tone-specific f0 downtrends.

The proposed dynamical f0 hard-landing model includes two global intonational units for H and L, and one dynamical boundary L% unit. The global H and L units are two free-fall-style dynamical systems with three parameters each—initial height, initial velocity and acceleration. The L% has a blending mechanism, where the final f0 is the average of the f0 of L% and the original f0 of their own tone-specific dynamics. The fitted results are consistent with the observed patterns of initial f0 rising, final f0 hard-landing, and the tone-specific downtrends (See Fig.2).

The current work offers a dynamical treatment of intonational events, which can be incorporated into the prosodic planning dynamics component of Articulatory Phonology [2], [3]. The goal of

individual lexical tones or pitch accent tones is hypothesized to be modulated by the proposed intonational dynamical units like global H/L downtrend units and final L%. Moreover, the current modelling work reconceptualizes the proposal of 'soft landing' versus 'hard landing' by Laniran & Clements (2003) in the framework of dynamical systems. A fuller picture of the dynamics of intonational units and their coordination patterns across languages remain to be investigated by future studies.

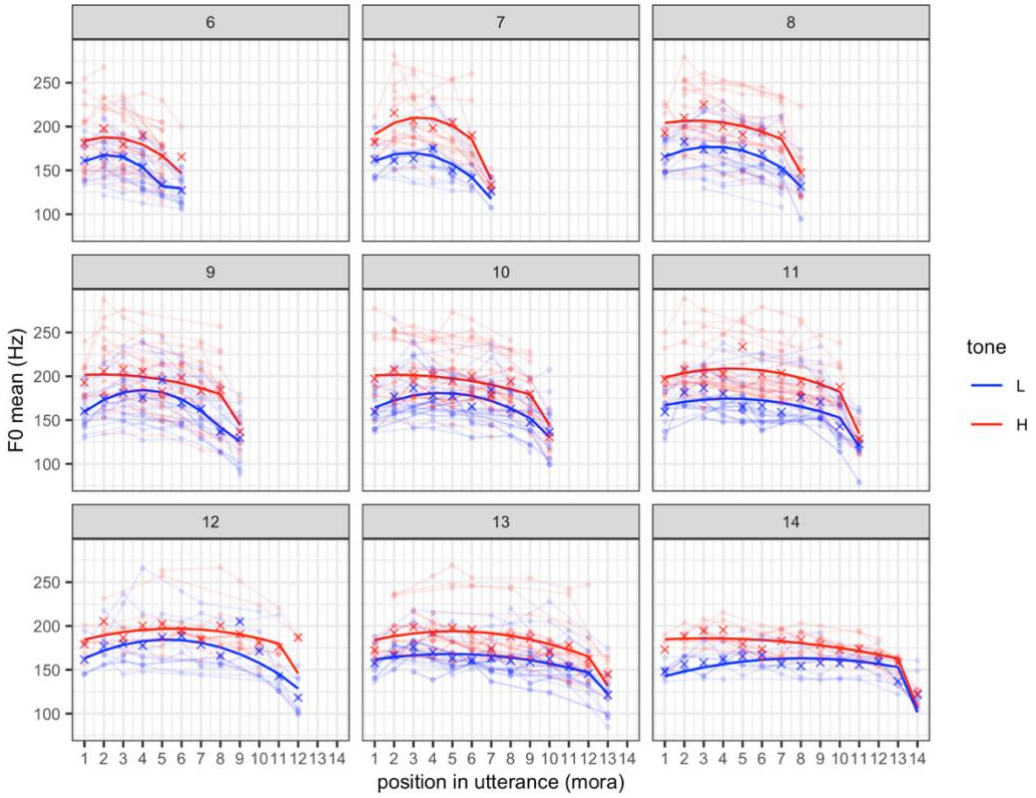

Fig.1 The fitted results based on the model versus the real data. The solid line—model predictions; The cross—averaged raw f0 values; The transparent dots and lines—individual data points.

[1] K. Iskarous, "The relation between the continuous and the discrete: A note on the first principles of speech dynamics," *J. Phon.*, vol. 64, pp. 8–20, 2017, doi: 10.1016/j.wocn.2017.05.003.

[2] E. Saltzman and K. Munhall, "A dynamical approach to gestural patterning in speech production," *Ecol. Psychol.*, vol. 1, no. 4, pp. 333–382, 1989, doi: 10.1207/s15326969eco0104_2.

[3] C. P. Browman and L. M. Goldstein, "Towards an articulatory phonology," *Phonol. Yearb.*, vol. 3, pp. 219–252, 1986, doi: 10.1017/s0952675700000658.

[4] D. Byrd and J. Krivokapić, "Cracking prosody in articulatory phonology," *Annual Review of Linguistics*, vol. 7. Annual Reviews Inc., pp. 31–53, Jan. 04, 2021. doi: 10.1146/annurev-linguistics-030920-050033.

[5] L. J. Downing and R. Rialland, "Introduction," in *Intonation in African Tone Languages*, 2016, pp. 1–16. doi: 10.1515/9783110503524.

[6] M. Y. Liberman and J. B. Pierrehumbert, "Intonational invariance under changes in pitch range and length," in *Language Sound Structure*, 1984, pp. 157–233.

[7] Y. O. Laniran and G. N. Clements, "Downstep and high raising: Interacting factors in Yoruba tone production," *Journal of Phonetics*, vol. 31, no. 2. pp. 203–250, 2003. doi: 10.1016/S0095-4470(02)00098-0.

[8] A. Rialland and M. E. Aborobongui, "How intonations interact with tones in Embosi (Bantu C25), a two-tone language without downdrift," in *Intonation in African Tone Languages*, 2016, pp. 195–222. doi: 10.1515/9783110503524-007.

[9] A. Rialland *et al.*, "Parallel corpora in Mboshi (Bantu C25, Congo-Brazzaville)," in *LREC 2018 - 11th International Conference on Language Resources and Evaluation*, 2019, pp. 4272–4276.