Distortion in Tone Production due to the Lombard Effect

Giang Le[1] Chilin Shih[1, 2] Yan Tang[1, 3]

[1]*Department of Linguistics, University of Illinois Urbana Champaign,* [2] *Department of East Asian Languages and Cultures, University of Illinois Urbana Champaign,* [3] *Beckman Institute for Advanced Science and Technology, University of Illinois Urbana Champaign*

In adverse listening conditions, human talkers spontaneously adapt the way they speak to maximize the success of communication due to the Lombard effect. Lombard speech is observed to have many changes at acoustic level over normal speech produced in quiet, such as raised F0 and overall intensity, reduced spectral tilt, expanded vowel space as well as elongated vowel duration (e.g. [1, 2, 3]). Some of the changes lead to better speech intelligibility even for non-native speakers [4]. At suprasegmental level, tone production in Cantonese has also been found to be influenced by the Lombard effect and produced with higher F0 [5]. F0 and F0 contour in Thai lexical tones are also seen to be emphasized in noise [6]. With the elevation of the overall F0 floor, one aspect less clear is whether the dynamic range of F0 contour would be reduced when the Lombard effect and hyper-articulation reach the ceiling, especially for tones realized by F0 fluctuations, such as rising and falling-rising tones. Consequently, confusions between tones could occur under a strong Lombard effect.

This study investigates the Lombard effect of different levels on the tone articulations in Northern Vietnamese. Northern Vietnamese presents an especially interesting case for the study of Lombard speech, for not only are its tones contrastive in F0 as is normally expected of tones, but also in voice quality. In studying this variety, we are interested in examining changes in the acoustic and laryngograph signals produced in noise, centered primarily on parameters associated with tone and sonority. Our inquiry and research design are specifically motivated by the following research questions: 1. How is the dynamic range of F0 contour of the lexical tones in Northern Vietnamese affected by hyper-articulation to different extents? 2. What impact does noise level have on the realizations of different phonation types in Northern Vietnamese. We hypothesize that due to the limited capacity of hyper-articulation of human talker, the dynamic range of F0 contour may be decreased, potentially resulting in reduced distinction between lexical tones. Furthermore, we anticipate that voice quality and phonation characteristics associated with the various lexical tones would also see changes, for example creaky tones would have increased regularity and harmonics-to-noise ratio at a higher noise level due to the extra vocal effort made to produce Lombard speech. This could potentially make tones that only differ in phonation type in some languages less distinguishable.

In the experiment, the target words in isolation and in carrier sentences (e.g., 'Tôi nói cho bạn nghe X bây giờ' meaning 'I speak X to you now') produced by several female and male talkers of the Northern Vietnamese variety are being recorded. The target words are chosen from representative syllable types (V, CV and CVC), with three vowels /e/, /a:/ and /u/, combined with six possible phonemic lexical tones listed in Table 1. The speech is produced in three conditions: in quiet as the control condition, and in the presence of speech-shaped noise at 78 and 90 dB SPL, which may lead to the Lombard effect in different degrees. Both acoustic and laryngographic signals are collected from the talkers.

The recordings from one male and one female talker were preliminarily analyzed. The interim results suggest that duration, intensity and mean F0 all increase in the noise conditions, echoing the previous findings on Lombard speech. Averaging across the two talkers, the time-normalized F0 contours [7] of the /a:/ sound with the B1 and C1 tone in the three noise conditions in Figure 1 show that the F0 contour are elevated, and that the dynamic range of the F0 contour is increased in the noise conditions ("90 dB" and "78 dB") compared with the "quiet" condition. However, "90 dB", which is expected to trigger a stronger Lombard effect, resulted in a smaller dynamic range than "78 dB". The vocal folds regularity of the creaky tones B2 and C2 tends to increase with a decreased jitter value and an increased HNR in noise, as indicated by the current data. The primitive evidence seems to be in favor of the hypotheses that hyper-articulation due to the Lombard effect may cause tone distortions during production. Further acoustic and laryngographic analyses will be performed to shed light on how F0 contour, phonation and laryngeal activities are affected when speaking in noise, hence distorted tone productions in Lombard speech. As an extension, if the distortions prove to be significant, further listening experiment is therefore warranted to inspect their impact on tone perception.

Table 1. Northern Vietnamese Phonemic Tones, F0 Contours and Phonation Characteristics [8]

Tone	F0 contour	Phonation characteristics
A1	level	modal voice
A2	low to mid falling	modal voice, or a lax breathy voice
B1	mid-rising	modal voice
B2	mid-falling	strong glottalization at the end
C1	falling tone with a similar F0 contour as tone A2, sometimes realized with a mid-falling-rising contour, more similar to the contour of C2	slight laryngealization at the end
C2	rising tone with glottal interrupt	creaky voice

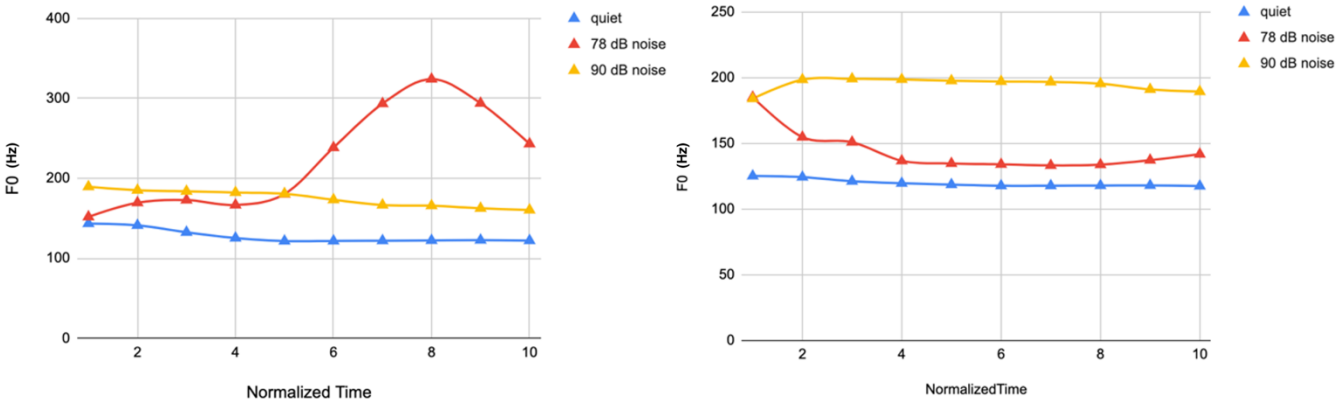

Figure 1. F0 vs. Normalized Time of vowel /a:/ in tone B1 (left) and tone C1 (right)

References

[1] Boril H. and Pollak P. (2005). Analysis of Lombard effect appearance in several Czech databases. In Electronic speech signal processing conference; 16th, Electronic speech signal processing.

[2] Garnier, M., Bailly, L., Dohen, M., Welby, P. and Loevenbruck, H. (2006): "An acoustic and articulatory study of Lombard speech: global effects on the utterance", In INTERSPEECH-2006, 1862

[3] Cooke, M. and Lu. Y. (2010). Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. J. Acoust. Soc. Am., 128(4): 2059–2069.

[4] Cooke, M., and Lecumberri, M. L. G. (2012). The intelligibility of lombard speech for non-native listeners. The Journal of the Acoustical Society of America 132 2, 1120–9.

[5] Zhao, Y. and Jurafsky, D. (2009). “The effect of lexical frequency and Lombard reflex on tone hyperarticulation.” J. Phonetics 37 (2009): 231-247.

[6] Kasisopa, B., Attina, V. and Burnham, D. (2014). The Lombard effect with Thai lexical tones: An acoustic analysis of articulatory modifications in noise. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. 1717-1721.

[7] Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France. 7-10.

[8] Nguyen, L. and Edmondson, J. (1997). Thanh dieu va chat giong trong tieng Viet hien dai (The tones and voice quality in modern Northern Vietnamese: instrumental case studies). Ngon Ngu (Linguistics), 1: 1-16.