EFFECT OF LEVEL AND TYPE OF NOISE ON FOCUS RELATED PROSODY

***Martti Vainio*[1], *Antti Suni*[1], *Anja Arnhold*[1,3], *Tuomo Raitio*[2], *Henri Seijo*[2], *Juhani Järvikivi*[4], *Daniel Aalto*[1,2], and *Paavo Alku*[2]**

[1]Institute of Behavioural Sciences (SigMe Group), University of Helsinki, Finland
[2]Department of Signal Processing and Acoustics, Aalto University, Finland
[3]Department of Cognitive Linguistics, Goethe-University Frankfurt am Main, Germany
[4]Language Acquisition and Language Processing Lab, NTNU, Norway

Introduction: Lombard effect

- Noise and hearing-loss induced reflex that makes people "raise their voice" [1].
- Has been attested in animals as well – thus partly a low-level biological phenomenon [2].
- Louder voice also raises the voice fundamental frequency (f_0).
- The reflex or effect is thought to involve both private and public feedback loops [3].
- Some, but not much is known about how it affects focus marking with f_0 [4].

Research questions

- How does noise affect f_0 contours in general? Apart from raised pitch, are there further changes?
- Are there differences between types of noise and noise levels regardless of equal loudness?
- Do the linguistically motivated f_0 features remain the same as in speech without the Lombard effect?

Materials and Methods

- 21 people (11 female) were recorded in an anechoic chamber.
- The participants read aloud three different types of sentences with differing focus conditions as responses to questions.
- The material was recorded with three different types of noises played to the participants at three different levels using headphones.
- Subjects' own voice was also played back via the headphones at constant level between subjects.

Noise types

- Three type of noises were used: Babble noise, Low-pass (white) noise under 1kHz, and White noise
- Three levels (plus no noise) were used.
- The noises were scaled to have equal loudness (ANSI S3.4-2007) at 4.75, 9.5, and 19 sones corresponding roughly to 60, 70, and 80 dB(A) sound pressure levels.

- Subjects were recorded with a calibrated high-quality condenser microphone in an anechoic chamber.

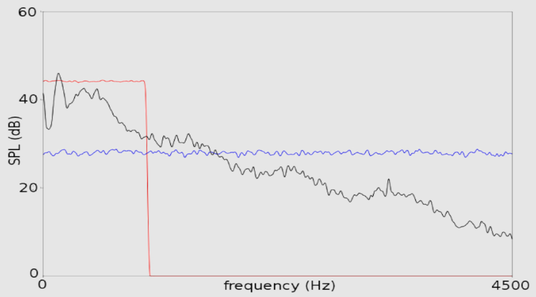

Figure 1: *Representative spectra of the three different noise types used in the study; red = low-pass noise, black = babble noise, blue = white noise. The figure shows the spectra between 0 and 4.5 kHz.*

Linguistic factors

- Sentences with three different focus conditions were elicited: *broad focus*, *narrow focus on subject* (the first word) and *narrow focus on object* (the last word).
- The f_0 contours for these conditions are well known and we expected them to remain unaltered [5, 6].
- The utterances were of the form e.g., *Paavi tavaa suuraa* (The pope reads a sura), and long vowels [ɑ], [i], and [u] were used in the subjects and objects.

Results

- All f_0 values were transformed to semitones (re 100Hz).
- Mean f_0 was calculated from nine points of interest (three per word).
- The f_0 expansion was calculated from the same points of interest based on the Bounded Variation (BV) norm:

$$\text{Expansion}(f_0) = \int_{T_{beg}}^{T_{end}} \left| \frac{df_0(s)}{dt} \right| ds$$

where $f_0(t)$ is the fundamental frequency at a given time point and T_{beg} and T_{end} are the beginning and end times of the utterance.

- The BV norm captures the overall movement in the contour in a time-invariant manner

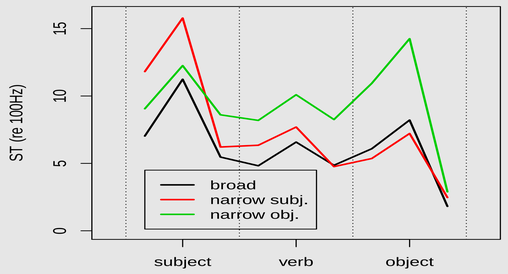

Figure 2: *Average f_0 contours calculated from three f_0 points per each word for different focus conditions for all noise types and levels. Black = broad focus, red = narrow focus on subject, green = narrow focus on object.*

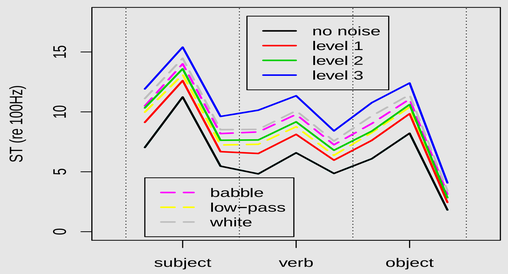

Figure 3: *Average f_0 contours for different noise levels and types.*

Statistical analyses (linear mixed-effects models) show that:

- **f_0 mean:**
  - The noise levels differ significantly.
  - The low-pass noise has a significantly lower mean f_0.
  - The different different focus conditions are also different from each other: i.e, the mean f_0 is generally lower when the narrow focus occurs on the first word (N1) and higher when it occurs on the last word (N2).
  - There is also a significant low-pass-noise:noiselevel3 interaction showing that the f_0 level is increased less in high level low-pass noise.
  - The gender difference is highly significant with the males speaking almost an octave lower than the females.
- **f_0 expansion:**
  - The contours are significantly influenced by the focus condition as well as noise levels.
  - The low-pass noise, however, does not differ from babble noise, but the contours are again more expanded in white noise.
  - There are also a white-noise:level interactions.
  - There are no significant gender differences.

Discussion

- As expected the f_0 level rises as a function of noise level. However, there are also differences between noise types with the low-pass noise affecting the levels less and white noise more, regardless of equal loudness.
- In addition to the typical f_0 level increase, there is an exponential increase in f_0 movement when the noise level increases. The effect is similar to the f_0 level with regard to different noise types.
- The fact that the noise types have different effects on f_0 points to a specific importance of pitch in auditory feedback.

References

[1] E. Lombard. Le signe de l'elevation de la voix. *Ann. Maladies Oreille, Larynx, Nez, Pharynx*, 37(101-119):25, 1911.

[2] J.W. Bradbury and S.L. Vehrencamp. Principles of animal communication. *Behavioral Ecology*, 12:283–286, 1998.

[3] H. Lane and B. Tranel. The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 14(4):677, 1971.

[4] R. Patel and K.W. Schell. The influence of linguistic content on the Lombard effect. *Journal of Speech, Language, and Hearing Research*, 51(1):209, 2008.

[5] Martti Vainio and Juhani Järvikivi. Focus in production: Tonal shape, intensity and word order. *The Journal of the Acoustical Society of America*, 121(2):EL55–EL61, 2007.

[6] Anja Arnhold. Multiple prosodic parameters signaling information structure: Parallel focus marking in finnish. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS), Hong Kong*, 2011.