Fundamental Frequency Histograms Measured by Electroglottography During Speech: A Pilot Study for Standardization

*†Romain E. Kania, *†Dana M. Hartl, *†Stéphane Hans, ‡Shinji Maeda, †Jacqueline Vaissiere, and *†Daniel F. Brasnu

Paris, France

Summary: This study was designed to develop a database for the electroglottographic measurement of fundamental frequency (Fo) in normal subjects in running speech, for reference in the diagnosis and follow-up of dysphonic patients. A prospective pilot study included 20 healthy male volunteers without laryngeal disorder. Electroglottographic recordings of speaking Fo during connected speech (French) were obtained from two texts with different prosodic content. Fo histograms were sensitive to the variation of speaking Fo between both texts. Graphic representation of the range and distribution of the Fo of the speaker were designed as normalized Fo histograms with plot lines at 5th and 95th percentiles. Less than 5% variability of Fo histograms was recorded when recording more than 15 subjects. This pilot study designed a graphic display of standardized electroglottographic Fo measurements during the physiological condition of connected speech. As the degree of Fo variability depends on the phonetic contents of the text and on the language spoken, a separate histogram for normal subjects needs to be developed in each country or at least for each voice laboratory, with a standard, previously chosen text.

Key Words: Fundamental frequency—Electroglottography—Standardization—Diagnosis—Speech.

INTRODUCTION

The fundamental frequency (Fo), which is defined as the number of vocal fold contacts during 1 second, reflects several laryngeal and extralaryngeal characteristics of a speaker, such as laryngeal size, subglottic pressure, and voicing mode. Fo varies during connected speech in that it is also affected by the phonetic and the prosodic contents of speech.

Electroglottography (EGG) is an interesting modality for measuring Fo.[1,2] EGG is a noninvasive method, which is based on Ohm's law that measures

Accepted for publication January 11, 2005.

Presented at the Annual Meeting of the American Academy-Head and Neck Surgery Society, San Diego, CA, September 23, 2002.

From the *Voice, Biomaterials and Head and Neck Oncology Research Laboratory, University Paris V, Department of Otorhinolaryngology-Head and Neck Surgery, Georges Pompidou European Hospital, Paris, France; †Institut de Linguistique et de Phonétique Générales et Appliquées, Université de la Sorbonne Nouvelle-Paris III, Paris, France; ‡ENST (Ecole Nationale Supérieure des Télécommunications), Paris, France.

Address correspondence and reprint requests to Daniel F. Brasnu, MD, Department of Otorhinolaryngology-Head and Neck Surgery, Hôpital Européen Georges Pompidou 20, rue Leblanc, 75015 Paris, France. E-mail: daniel.brasnu@hop.egp.ap-hop-paris.fr

Journal of Voice, Vol. 20, No. 1, pp. 18–24
0892-1997/$32.00

doi:10.1016/j.jvoice.2005.01.004

the variation in electrical resistance between two electrodes placed on each side of the thyroid cartilage.[3,4] The electroglottographic wave reflects variations of the vocal fold contact area.[4–13] EGG provides a physiological measure of Fo at the level of the laryngeal source.[14,15] The EGG signal is much easier to process in comparison with the acoustic signal.[14]

An advantage of EGG is that it can measure Fo during connected speech because no measurement sensors are introduced into the mouth. Whereas sustained vowels are often the basis for objective evaluation, Fourcin pointed out the interest of assessing connected "real-life" speech.[16] Fourcin[3] first described histograms of Fo during connected speech as a representation of the range and distribution of the vocal fold frequencies.[6]

The issue remains that no standard Fo histograms are usable as a normalized database. The absence of such normative values limits its practical application.[14] As long as normal values have not been validated, the clinical application of EGG will be limited.

The aim of this study was to determine the speech pattern of normal electroglottographic values for Fo histograms in French.

SUBJECTS AND METHODS

Subjects

Twenty healthy male volunteers were prospectively studied. All subjects gave informed consent to participate in this study and underwent medical history and laryngeal examination. Inclusion criteria were male subjects, native French speakers, aged between 25 and 45 years, with no history of laryngeal or neurologic disease, no complaint of voice or speech deterioration, and a laryngeal examination with normal findings. Ages ranged from 26 to 39 years (mean, 32.5 years). To generalize the measurement for routine practical application, we included patients having had past light tobacco intake or very mild current consumption. Thirteen subjects were nonsmokers, and 7 had tobacco intake of less than one pack of cigarettes per day. None were professional singers. All subjects were evaluated and recorded by the same author (R.E.K).

Instrumentation

The EGG Laryngograph (Laryngograph Ltd, London, U.K.) measured the EGG signals that were displayed and analyzed by a computer data-processor SESANE (SQLab, Aix-en-Provence, France). The EGG device recorded the electric impedance across the neck between two electrodes placed on either side of the thyroid cartilage and held in contact with the skin by an elastic collar. The output of the EGG device was processed by an electronic preamplifier (DIANA; SQLab, Aix-en-Provence, France) and then by a 16-bit analog-to-digital (A/D) converter that was included in a 366-MHz Pentium personal computer (Gateway 2000; Gateway, Irvine, CA). The *SESANE* (SQLab, Aix-en-Provence, France) software was simultaneously applied for acoustic signal acquisition and analysis. An AKG (Vienna, Austria) model C410 condensor microphone mounted on a headset was connected to the personal computer via the A/D converter DIANA for acoustic recordings. Both electroglottographic and acoustic measures were displayed by SESANE software.

Procedure

The following procedure was implemented for each recording. The microphone was placed on the side of the mouth at 8 cm from the labial commissure. The sound intensity was calibrated with a buzzer placed 22 cm in front of the microphone before each recording. The subjects were seated upright. Electrodes of the Laryngograph were placed approximately 1.5 cm laterally to the anterior angle of the thyroid cartilage with no conducting gel between the electrode and the skin. The subject was instructed to speak normally at a comfortable pitch and intensity level. The connected speech task consisted in the reading of two texts at a comfortable pitch and intensity. Text 1 was a story, and text 2 included 20 sentences (see the Appendix) from a database statistically representative of the distribution of phonemes in the French language (CNET, technical note 1980, NT/LAA/TSS/26). One single trial of each text was obtained. The reading of each text lasted more than 1 minute.

Data analysis

The simultaneously recorded EGG and acoustic signals were directly digitized. EGG signals were

only applied for Fo quantification. The EGG signal was filtered with a fourth-order Bessel filter to delete nonphonated intervals for analysis. The experimenter checked on the screen that the Fo display was synchronous with segments of voicing. A peak-to-peak detection algorithm was applied for Fo measure. The upper portion of the EGG wave was considered and configured as the maximal glottal closure, and the lower portion of the EGG wave was considered and configured as the maximal glottal opening. Results were expressed for mean Fo, standard deviation, variance, and range for each task performed. The graphic representation of the range and distribution of the Fo of the speaker by Fo histograms was normalized as plot lines of percentiles. We also studied the variation between Fo results of text 1 and text 2 to choose the most suitable one for more application of the standardization.

Difference of mean Fo histogram area ΔA according to the number of subjects

One aim of this pilot study was to approximate the required number of subjects to establish the normalization per age group. Therefore, the variability of Fo histograms was studied according to the cumulative number of subjects recorded. Each mean distribution histogram was displayed after inclusion of one subject after another by random order. A new histogram was calculated as the difference between the mean Fo histogram including n subjects and the mean Fo histogram for n-1 subjects. The area difference between successive histograms, ΔA, was plotted as a function of the number of subjects included.

Statistical analysis

Statistical analysis was performed with the *StatView* software (SAS Institute Inc., Cary, NC). Descriptive statistics included mean, standard deviation, variance, coefficient of variation, minimum, maximum, and range. Repeated analysis of variance (ANOVA) compared the means and distribution of Fo between (1) both texts and (2) between smokers and nonsmokers. Plot lines at 5th and 95th percentiles became the graphic representation of Fo histograms. A graphic method based on histograms and plots defined the normality of the mean distribution of Fo histograms for this group of normal subjects. Statistical significance was set at the level of 0.05.

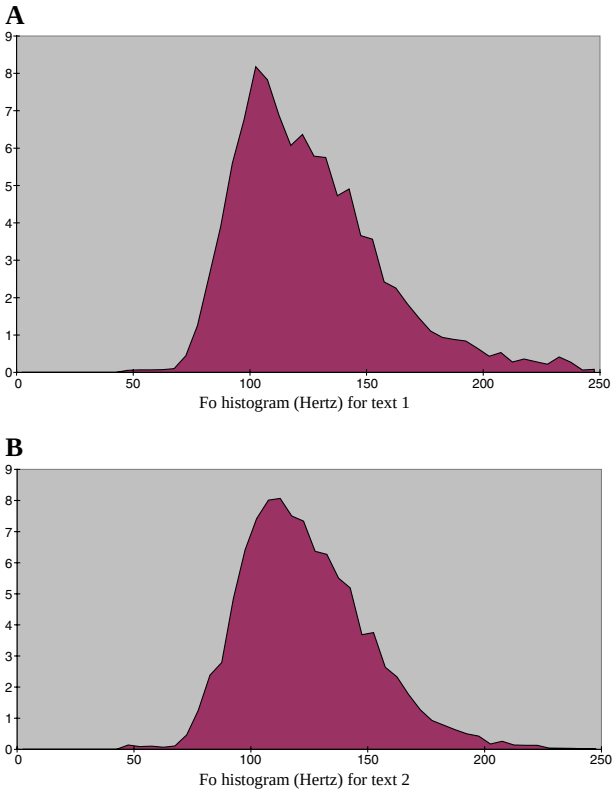

FIGURE 1. Fo histograms for text 1 (**A**) and text 2 (**B**) display the percentage (Y axis) of each Fo value (X axis) over the total duration of the voiced text.

RESULTS

Fo histograms

Figure 1 illustrates the mean of the Fo distributions across subjects for each text. Graphic representations with mean and plot lines at 5th and 95th percentiles for each text are given in Figure 2. The histograms presented in Figures 1 and 2 showed apparent skewing to the right. Descriptive statistics concerning Fo for each text are summarized in Table 1. The measurements of dispersion (standard deviation, variance, range, coefficient of variation, coefficient of asymmetry) were higher in text 1 than those in text 2, but the differences were not statistically significant. There was neither a significant difference between smokers and nonsmokers nor a significant

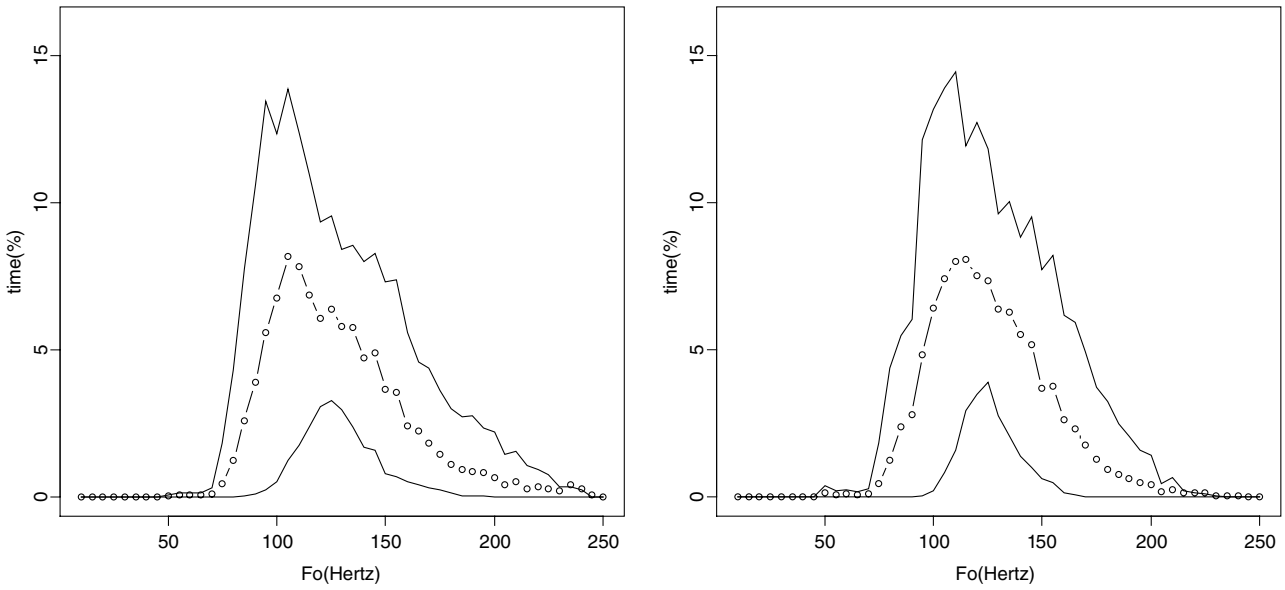

FIGURE 2. Graphic representations of Fo histogram with mean (dotted line) and plot lines at 5th (inferior curve) and 95th (superior curve) percentiles for both texts.

effect of the day of recording on the mean value of Fo.

Difference of mean Fo histogram area ΔA according to the number of subjects

Figure 3 illustrates the drop in ΔA when subjects were added. Such a drop means that the more subjects are included, the less the mean distribution of Fo histogram varies. After inclusion of 15 subjects, no more than 5% variability was recorded when adding a new subject in this definite age group.

DISCUSSION

The need for standardization of procedures of voice measurement has been emphasized for both clinical and research applications.[14,17] The main interest of this pilot study was to design a standardized graphic representation of speaking Fo during connected speech with EGG.

EGG is a method recording Fo directly at the level of the laryngeal source.[15] Therefore, the Fo measurement is not affected by the supraglottic resonances occurring in the vocal tract. EGG determines Fo based on laryngeal physiology and is thus, in theory, not affected by the random noise prevalent in dysphonic voices, as compared with Fo extraction with acoustic algorithms.[18–20] The Fo distribution calculated from running speech would seem to better reflect the habitual voice use of the speaker, as compared with pitch determination from a sustained vowel, which is in most cases for non-voice professionals an "unnatural" vocal gesture. The phonation of a sustained vowel explores different aspects of the phonatory function such as the long-term stability of Fo for the evaluation of voice quality. We agree with Fourcin,[16] who emphasized the virtue of connected speech as it pertains to the physiological aspect of voice and speech assessment as a whole. We think that the connected speech is appropriate for clinical and physiological evaluation of voice.

Fundamental frequency histograms is perhaps the simplest way to analyze speaking Fo because it is nothing more than a bar graph of the occurrences of Fo in the sample.[2] Although this form of analysis was thought to be one of the most useful information in voice evaluation,[21–24] the method is not currently in general application, probably because of the absence of rigorous standardization of fundamental frequency histograms.

TABLE 1. *Descriptive Statistics for Each Text*

	Text 1	Text 2
Fo in Hertz (mean ± SD)	130.5 ± 21.3	128.6 ± 17.9
Variance	452.1	319.4
Minimum (mean ± SD)	104.3 ± 12.1	100.3 ± 16
Maximum (mean ± SD)	177.5 ± 20.3	163.9 ± 25.4
Range	73.2	63.6
Median	124.9	128.1
Asymmetry	0.69	0.32

This lack of scientific data was the initial impetus for the current work based on a standardized procedure to obtain Fo histograms in a subgroup of male patients studied as a pilot study.

Our study demonstrated that Fo histograms were sensitive to the variation of Fo during connected speech. Higher measures of dispersion in text 1 might be from the normal variations in Fo during speaking, which were higher in text 1 than in text 2. Because text 1 is a popular story, its reading may be much more prone to prosodic variations.[25] The prosodic content of text 2 is globally neutral because no meaning results from the connection of the sentences. There is a lack of consensus regarding which text is better suited for routine practical purposes. We propose to choose one, which is the most representative of the phonemes of the language applied (text 2). In standardization, the same text, with the same punctuation, must be employed to apply the standardized plots as references. One limit of this study was that all results were obtained from a series of French-speaking subjects. Different languages may also demonstrate different Fo distributions. More studies are required to determine standardized plots for other languages.

The basic purpose of any standardization is to provide normative data about the mean and the variability of the measurements between persons. The issue of the Gaussian character of our Fo histogram is not the purpose of this study. In this pilot study, we were interested in elaborating the normalization of Fo histograms, ie, standardization with data from subjects considered as normal, for more application of EGG. We assessed normality for representative, reliable, valid, interpretable, generalizeable, and usable standards. Our concern as investigators was to design a standardization that would be suited for routine clinical purposes. The practical application of Fo histograms is evident because the variability (range and distribution) of speaking Fo in physiological conditions of connected speech is graphically displayed. For practical application, no statistical test is required. Any new recorded Fo histogram with the same text can be easily and directly compared with the graphic representation of plots lines of percentiles of normal subjects. A lot of commercially available software provides graphic display of results. Some developments may include standardization to compare Fo histograms recorded with that of normal subjects of the same sex and age. More studies are necessary to determine if a standard Fo distribution exists specific to each voice pathology that would aid in the diagnosis and follow-up of these disorders.

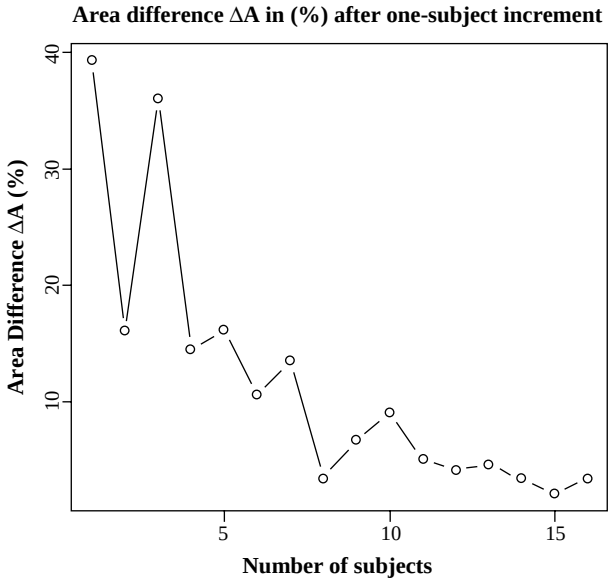

FIGURE 3. Difference of mean Fo histogram area ΔA in percent (Y axis) according the number of subjects included (X axis).

In this pilot study, we wanted to determine the number of subjects required for each age group to achieve a representative standardized curve, which may be assessed by ΔA. In our opinion, the number of subjects in our pilot study was sufficient for this age group because the inclusion of an additional subject beyond the 15th resulted in a variation not exceeding 5%. We can assume that this may be close to the number of subjects required for other age groups.

Acknowledgment: The authors thank Pierre Taupin, MD, and Paul Landais, MD, for their assistance in statistical analysis of the data.

REFERENCES

1. Askenfelt A, Gauffin J, Sundberg J, Kitzing P. A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. *J Speech Hear Res*. 1980;23:258–273.
2. Baken R. *Clinical Measurements of Speech and Voice. Vol. Vocal Fundamental Frequency.* San Diego, CA: Singular Publishing Group; 1987.
3. Fourcin AJ, Abberton E. First applications of a new laryngograph. *Med Biol Illus*. 1971;21:172–182.
4. Baken RJ. Electroglottography. *J Voice*. 1992;6:98–110.
5. Fourcin AJ. *Laryngographic Examination of Vocal Fold Vibration. Ventilatory and Phonatory Control Systems*. London: Oxford University Press; 1974:315–333.
6. Fourcin AJ. Laryngographic assessment of phonatory function. In: Ludlow CL, Hart MO, editors. *Proceedings of the Conference of the Assessment of Vocal Pathology.* Vol. 11. Rockville, MD: American Speech-Language-Hearing Association; 1981:116–127.
7. Baer T, Lofqvist A, McGarr NS. Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. *J Acoust Soc Am*. 1983;73:1304–1308.
8. Childers DG, Hicks DM, Moore GP, Alsaka YA. A model for vocal fold vibratory motion, contact area, and the electroglottogram. *J Acoust Soc Am*. 1986;80:1309–1320.
9. Gilbert HR, Potter CR, Hoodin R. Laryngograph as a measure of vocal fold contact area. *J Speech Hear Res*. 1984; 27:178–182.
10. Lecluse FLE, Brocaar MP, Verschuure J. The electrography and its relation to glottal activity. *Folia Phoniatr.* 1975;27: 215–224.
11. Rothenberg M. Some relations between glottal airflow and vocal fold contact area. In: Ludlow CL, Hart MO, editors. *Proceedings of the Conference on the Assessment of Vocal Pathology.* Vol. 11. Rockville, MD: American Speech-Language-Hearing Association; 1981:88–96.
12. Childers DG, Krishnamurthy AK. A critical review of electroglottography. *CRC Crit Rev Biomed Eng*. 1985;12: 131–161.
13. Orlikoff RF. Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. *J Speech Hear Res*. 1991;34:1066–1072.
14. Houben GB, Buekers R, Kingma H. Characterization of the electroglottographic waveform: a primary study to investigate vocal fold functioning. *Folia Phoniatr.* 1992; 44:269–281.
15. Kent RD, Read C. *The Source-Filter Theory of Vowel Production. The Acoustic Analysis of Speech*. San Diego, CA: Singular Publishing Group; 1992:18–21.
16. Fourcin A. Voice quality and electrolaryngography. In: Kent R, Ball M, editors. *Voice Quality Measurement*. San Diego, CA: Singular Publishing Group; 2000.
17. Hicks D. Functional voice assessment: what to measure and why. *NIDCD Monogr.* 1991;1:204–209.
18. Hess W, Indefrey H. Accurate time-domain pitch determination of speech signals by means of a laryngograph. *Speech Commun*. 1987;6:55–68.
19. Horiguchi S, Haji T, Baer T, Gould WJ. Comparison of electroglottographic and acoustic waveform perturbation measures. In: Baer T, Sasaki C, Harris KS, editors. *Laryngeal Function in Phonation and Respiration*. San Diego, CA: College Hill Press; 1987:509–518.
20. Haji T, Horiguchi S, Baer T, Gould W. Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. *J Acoust Soc Am*. 1986;80:58–62.
21. Pegoraro Krook MI. Speaking fundamental frequency characteristics of normal Swekish subjects obtained by glottal frequency analysis. *Folia Phoniatr.* 1988;40:82–90.
22. Pedersen MF, Moller S, Krabbe S, Bennett P. Fundamental voice frequency measured by electroglottography during continuous speech. A new exact secondary sex characteristic in boys in puberty. *Int J Pediatr Otorhinolaryngol*. 1986;11:21–27.
23. Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngol*. 1980;90:441–451.
24. Fourcin A, Abberton E, Miller D, Howells D. Laryngograph: speech pattern element tools for therapy, training and assessment. *Eur J Disord Commun*. 1995;30:101–115.
25. Monnot M, Orbelo D, Riccardo L, Sikka S, Rossa E. Acoustic analyses support subjective judgments of vocal emotion. *Ann NY Acad Sci*. 2003;1000:288–292.

APPENDIX

Text 1

Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon : un beau matin, elles cassaient leur corde, s'en allaient dans la montagne et là haut le loup les mangeait. Ni les caresses de leur maître, ni la peur du loup, rien ne les retenait. C'était, paraît-il, des chèvres indépendantes, voulant à tout prix le grand air et la liberté.

Le brave Monsieur Seguin, qui ne comprenait rien au caractère de ses bêtes, était consterné. Il disait : c'est fini ; les chèvres s'ennuient chez moi, je n'en garderai pas une.

Cependant, il ne se découragea pas, et, après avoir perdu six chèvres de la même manière, il en acheta une septième ; seulement, cette fois, il eut soin de

la prendre toute jeune, pour qu'elle s'habituât mieux à rester chez lui.

Ah ! qu'elle était jolie, la petite chèvre de Monsieur Seguin ! qu'elle était jolie avec ses yeux doux, sa barbiche de sous-officier, ses sabots noirs et luisants, ses cornes zébrées et ses longs poils blancs qui lui faisaient une houppelande !

Text 2

Leur chienne a hurlé toute la nuit.
Pour se protéger, il s'est couché près de ma porte.
Sa voisine est inimitable.
Le renard se hâte vers son gîte.
Le bouillon fume dans les assiettes.
Le caractère de cette femme est moins calme.
Le camp d'été s'est passé au bord du fleuve.
Un train entre déjà en gare.
Souvent je m'accoude au muret de ce pont.
A l'ouest, mes pommiers donnent peu.
Je ménage une surprise à mon ami.
Les boulangers façonnent des pains.
Vos livres devront être couverts.
Mangeras-tu de cette tarte aux prunes?
Le chapeau de Monique est sur la table.
Il s'est glissé loin des spectateurs.
Il s'empresse de réclamer ce qu'on lui a promis.
Vous tremblez parce que vous avez froid.
C'est le soir qu'il travaille le mieux.
Un serpent noir fuit sous une pierre.