Ann. Bull. RILP, No. 12, 113-125 (1978)

SOUNDSPECTROGRAPHIC ANALYSIS OF NARON CLICKS - A PRELIMINARY REPORT -

Ryohei Kagaya

1. Introduction

Click sounds characterized by the ingressive velaric airstream [1] are found in the Khoisan (the Hottentot and Bushman) languages, in some of the southern Bantu languages (e. g., Zule, Xhosa, etc.) and in the Sandawe and Hadza languages in Tanzania.

The articulation of a click sound may be generally sketched as follows [2] : Firstly, two closures are made in the vocal tract, the one by the contact between the blade of the tongue and the dental-alveolar or alveo-prepalatal region of the upper articulator (see below), and the other by the contact between the back of the tongue and the posterior part of the palate. The former closure is called the primary closure and the latter the secondary closure. In the southern dialects of the San (Bushman) languages, the primary closure is also made by the lips. Secondly, the center of the tongue which does not touch any parts of the palate is lowered and the air enclosed in the chamber formed by the two closures and the side-edges of the middle of the tongue is rarefied. When the place of the primary closure is released, the click noise is generated at the place of the primary closure by the air rushing into the chamber. Immediately after the release, the secondary closure is released and velar-frication, aspiration, ejective noise or glottalization may follow it.

In this study, the place where the first release takes place is called the place of primary release. In the case of the lateral affricate click (see below), the primary release is made at the side-edge(s) of the tongue.

Beach (1938) divides the phonetic characteristics of a click sound into the "influx" and the "efflux" from the view-point of the mechanisms of the airstream. He states as follows: "...the essential feature of a click is the "influx" of air into the mouth from without, in other words, the implosion. Clicks may therefore be classified in the first place according to the exact place of this influx and the exact manner in which it is made. But there is a second method of classification which must also be made, according to the "efflux" of air from the lungs. In Hottentot, the implosive part of the click is made by releasing the rim (of the back chamber) at some pre-velar point, while the back of the tongue remains against the velum. ...Other effects which may be produced by the "efflux" of air from the lungs are nasalization, voicing, and other modifications produced at the glottis. " (The words in parentheses are those of the author's.) That is, the influx means the click noise caused by the ingressive velaric airstream at the place of the primary release and the efflux signified the other phonetic characteristics involving the pulmonic egressive airstream and articulatory mechanism which are not required for the click noise production. In this study, the "influx" is called the "click (noise)" and the "efflux" is termed "click accompaniment according to Lanham's (1964) or Traill's terminology (1973).

Clicks have usually divided into five types according to both the place of primary closure and the manner of the release by many researchers. As to the place of primary closure of a few types, there may be dialectal or idiosyncratic differences among researchers' descriptions, but they are minor differences and may be neglected [3]. Here, the places may be divided into three regions along the upper articulator, i.e., the labial, the dental-alveolar and the alveo-prepalatal. The manners of primary release have also divided into two types by the duration of the click following the release. That is, one is characterised by a relatively long duration and the other by much less or none at all [4]. Symbols for the click are shown in Table 1.

Place of Primary Closure / Manner of Primary Release	labial	dental-alveolar	alveo-prepalatal
Short duration		≠	!
Long duration	⊙	/	//

Table 1. Symbols for five types of click noises. The place of the primary release is the same as that of the primary closure, except for // in which it is the side-edge(s) of the middle of the tongue.

2. Method of analysis

Frequency characteristics and duration of four types of click of Naron, i.e., /, ≠, // and ! , are studied by a soundspectrograph. On the soundspectrogram, a click can be characterized by the wide frequency range (almost 0 to 8 kHz) and the strong intensity, compared with other fricatives caused by the eggressive airstream. Along the frequency scale, the click is observationally divided into two parts, i.e., the strong intensity part and the weak one. In one case, the strong intensity part consists of two or more different frequency bands. Some soundspectrograms of clicks are examplified in Figure 1, where the horizontal line shows the duration and the vertical line(s) the strong intensity band(s). The frequency region(s) of the strong intensity band(s) of each click was measured in the soundspectrogram analyzed by a wide-band filter in the accuracy of 250 Hz (see Figure 2), where the strong intensity band changes temporally in frequency, the total frequency range is indicated in the figure. The duration is also measured in the accuracy of 5 msec (see Table 2).

Click Noise Type	/	≠	//	!
Duration of Click Noise (in msec)	30 (20 - 40)	10 (5 - 20)	30 (20 - 40)	10 (5 - 20)

Table 2. The average value of the duration of clicks. Each number in parentheses is the range.

All utterence samples are meaningful words and they are uttered in isolation by a male native speaker of the Naron (see below), In them, each click sound occurs word-initially. In many click sounds, the accompaniments are voiceless (plain) types. [6] In other cases, it may be any of six distinctive accompaniment types (Kagaya, 1978). Tones of click sounds are high-level or mid-level.

The speaker is Mr. Guek'ai /eni, aged 35, who was born and raised on a farm near Ghanzi, Botswana. He has never moved out of Ghanzi and speaks no other languages except for the Naron and Setswana, the latter of which is one of the southern Bantu languages.

3. Results of Soundspectrographic analysis

3-1. Dominant frequency region of Click noise

In the case of the / type, the number of samples is 38, and there is a clear tendency for the strong intensity part to appear above 3 kHz.

In the case of the // type, the number of samples is 34.. In 19 of them, the strong intensity part consists of two bands, the lower band of which appears in the frequency region between 0.5 kHz and 3.5 kHz, and the upper one in the region above 3.5 kHz. In the six samples in which the strong intensity part consists of more than two bands, the bands of each sample can be distributed into the two frequency regions mentioned above. In the nine samples, the strong intensity part consists of one band in the lower frequency region except in one. Even in the exceptional sample, the band does not exceed 4 kHz and the slightly strong band shown as a dotted line in the figure is observed below 3 kHz. Thus, it may be said soundspectrographically that the strong intensity part of // consists basically of two strong intensity bands and the main band is in the lower frequency region.

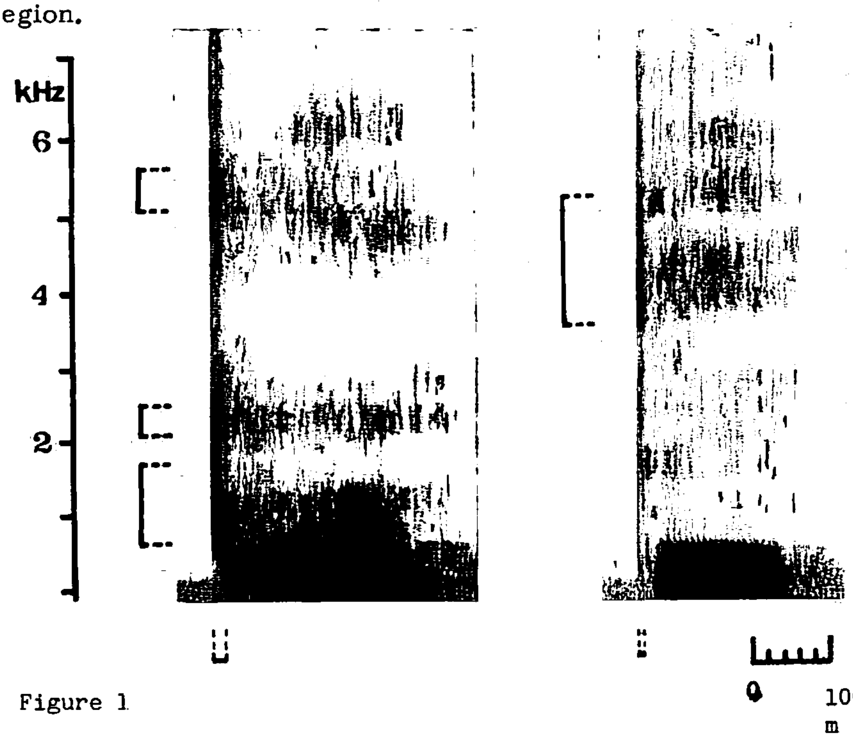

Fig. 1. Examples of soundspectrograms of [//õõ] (left: teeth) and [≠e] (right: ear). In the figure, the strong intensity part of //(left) consists of three bands, and that of ≠ of one band.

Figure 1

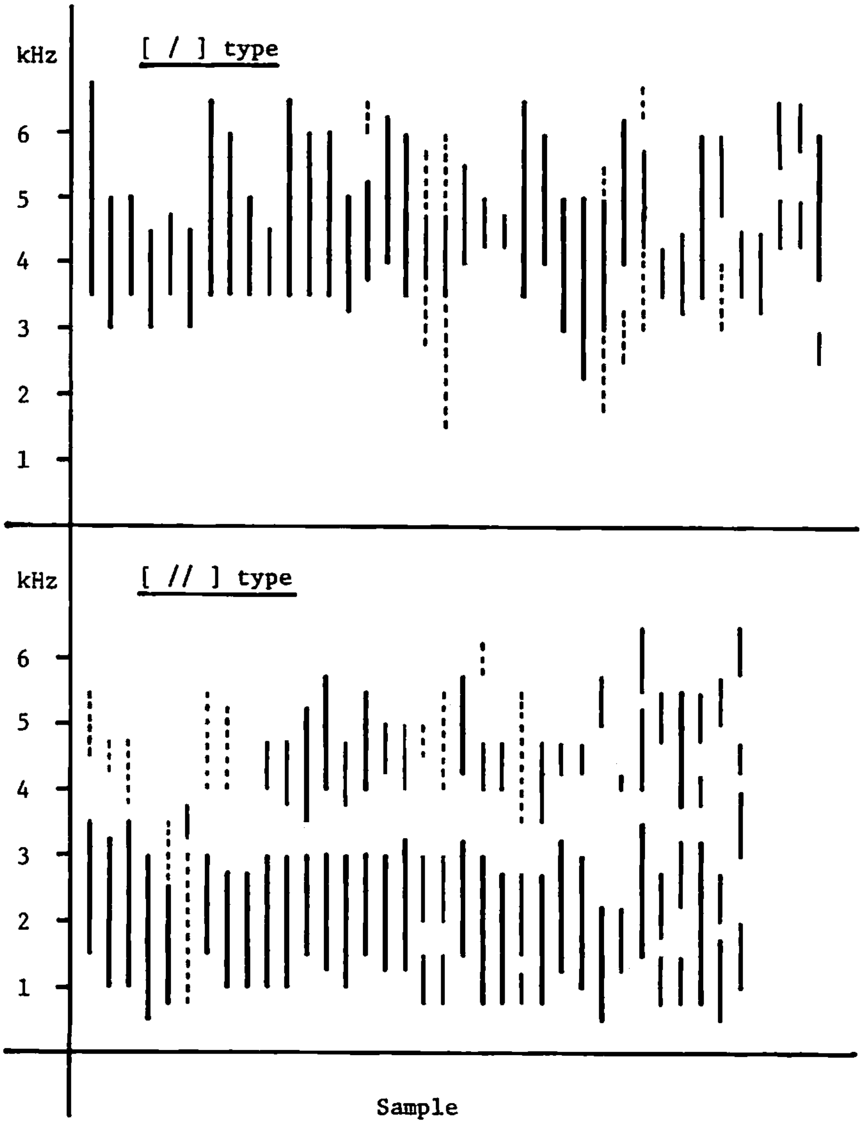

Figure 2a.

Fig. 2. a. The strong intensity part of each click noise type. The ordinate is the frequency scale in kHz, and samples are arranged on the abscissa. In each figure, the solid line shows the strong intensity part, and the dotted line the slightly strong intensity part. The weak intensity part is not shown in the figure.

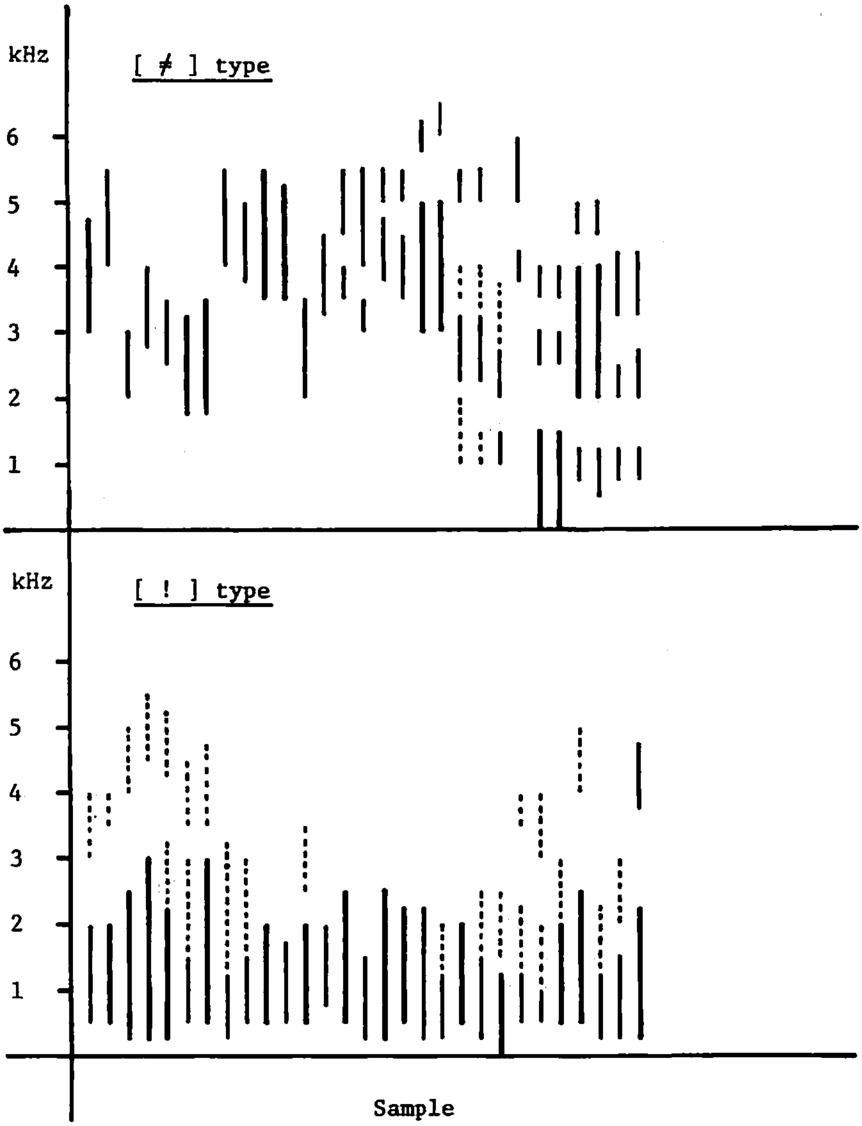

Figure 2b.

Fig. 2. b. The strong intensity part of each click noise type. The ordinate is the frequency scale in kHz, and samples are arranged on the abscissa. In each figure, the solid line shows the strong intensity part, and the dotted line the slightly strong intensity part. The weak intensity part is not shown in the figure.

In the case of the ! type, the number of samples is 29. In 28 of them, the strong intensity part consists of one band located in the frequency region between 0 and about 2.5 kHz. In one sample where the strong intensity part consists of two bands, the upper band is located above 3 kHz, but the lower one, the frequency width of which is wider than that of the upper one, is located in the region below 2.4 kHz.

In the case of the ≠ type, the number of samples is 29. The strong intensity part of 13 samples consists of one band located above approximately 2 kHz. In ten samples in which the strong intensity part consists of two bands, both are located above 2 kHz, except in one sample, where the main band with the wider frequency range is still in the same frequency region.

In the case of the 6 samples where the strong intensity part consists of three bands, the lowest band is below 2 Khz, but the two upper bands are in the frequency region above 2 kHz. Thus, it can be said that the dominant frequency of ≠ is located above 2 kHz.

From comparisons among the frequency characteristics of the four types of clicks, it can be said that (1) in a click type, the tendency of the frequency distribution of the strong intensity part is not influenced by either the click accompaniment or the tone; (2) the click noise intensities of both the / and ≠ types are dominant in the relatively high frequency region, while those of both the // and ! types in the low frequency region; and (3) the strong intensities of both the ≠ and // types are distributed in a wide frequency range compared with those of the / and ! types.

3-2. Duration of Click

The duration of each click type was measured on the soundspectrogram in the accuracy of 5 msec as shown in Table 2.

It is noted that (1) the influences on the duration by either the tone or the accompaniment of the click sound are not found. (2) The duration of the ≠ type and the ! type is 10 msec on the average and that of the / type and the // type 30 msec, respectively. (3) The range of the durational variation within the same type is very small, that is, there is little overlapping in the duration between the types with shorter noise (≠ and !) and those with longer noise (/ and //).

3-3. Other acoustic characteristics

In cases of both ≠ and ! , the temporal change of the frequency(-ies) of the strong intensity band(s) of the click noise cannot be observed on the soundspectrogram, since the duration is very short, but in many cases of both / and //, on the other hand, it is clearly observed that there is a downward frequency shift of the strong intensity band(s) with time.

4. Discussion

4.1. Characteristics of Resonant frequencies

In order to study the characteristics resonant frequencies of click noise, the shape of vocal tract during the click generation must be estimated. It may be plausible that the vocal tract configuration is roughly divided into three parts, i.e., the front cavity, the constriction and the back cavity. Furthermore, these lengths of the three parts during the click generation may be assumed to be the same as the length of the front cavity, that the the primary closure and that of the closed back chamber at the time immediately before the primary release, respectively.

In the following, these lengths shall be estimated based on both the static palatograms by Beach (1938) and some summarized data of the X-ray film by Traill (1976, personal talk), the latter of which were taken from the lateral direction, the frame rate being 50 frames/sec. The subject of the palatogram was of the Nama of Khoi (Hottentot) and that of the X-ray film of the !xõ of the southern group of the San (Bushman). Although these languages are different from the Naron analyzed here acoustically, it is assumed that there is no (or little) difference in both the articulation and the acoustic characteristics of the corresponding click type (click represented by the same symbol) among the three languages. The Nama has four types of clicks, i.e., / , ≠ , // and ! (Beach, 1938), as in the Naron (Bleek, 1928). The !xõ has a bilabial type besides the above four types (Traill, 1973.).

Neither Beach's static palatograms nor Traill's summarized data of X-ray film can show directly the shape of the vocal tract immediately before the primary release for the following reasons. In the case of the static palatograms, the contact pattern represents the total area where the tongue contact takes place throughout the time course of articulation, where the contact pattern varies from time to time. In fact, it is observed on the X-ray film that the contact area (of the lateral view) of the primary closure immediately before the release decreases considerably compared with the maximum contact area during the click articulation, with the volume expansion of the back chamber formed by both the primary closure and the secondary (velar) closure. Therefore, it is difficult to regard the contact pattern of the static palatogram as that immediately before the primary release. In the case of the summarized data of X-ray film, the temporal change in the size of the back chamber is clearly shown until the time immediately before the primary release, but the front point of the primary closure is not given on the summarized data.

The front point of the primary closure immediately before the release was estimated from the static palatogram under the assumption that the movement of the front point of the primary closure is little during the time course of articulation. Then, the length of the front cavity is estimated as the distance between the lip and the front point of the primary closure. The length of the back cavity is directly observed on the data immediately before the primary release of X-ray film. The length of the constriction is estimated as a distance between the back point of the front cavity and the front point of the back cavity.

In the case of the // type, the length of the front cavity may be longer than the distance between the lip and the back point of the primary closure, the latter of which is observed on the data of X-ray, since the place of the primary release is the side-edge(s) of the middle of the tongue. The length of the constriction may be shorter than that observed in the static palatogram, since the back cavity may be expanded before the primary release as in the cases of the other click types. Similarly, the length of the back cavity may be longer than the transverse length of the back cavity observed on the static palatogram.

According to these estimations, the length of each of the three parts, examined is summarized as follows (see also Table 3).

Click noise type	/	ǂ	//	!
Length of front cavity	short	short	long	middle
Length of constriction	short	long	short	short
Length of back cavity	very long	middle	middle	middle

Table 3 : Length of each cavity estimated from both Beach's static palatograms (1938) and Traill's X-ray films (1976, personal talk).

(1) The length of the front cavity for both the / and ǂ type is short compared with that for both the // and ! types. (2) The length of the primary closure (i. e., the constriction) is considerably greater compared with that in the other types. (3) The length of the back cavity for the / type is longer compared with those of the other types.

Fant (1960) calculated transfer functions of various vocal tract shapes of stop consonant at the time immediately after the release, i. e., 10 to 20 msec later than the release, and his results were in good agreement with the acoustic measurements of frequency characteristics of the consonant. These shapes are different from those of clicks in many points, e. g., in the absolute length or cross-sectional area of each cavity and in the exact location of the noise source, etc. But, at the time immediately after the release, the vocal tract shape of a stop consonant consists of three cavities, i. e., the front cavity, the constriction and the back cavity, as in the case of clicks.

From the data of Fant, it may be said that (1) the shorter the length of the front cavity, the higher the resonant frequencies, under the condition that the lengths of both the constriction and the whole vocal tract are constant. Thus, the resonant frequencies of the / type tend to be higher compared with those of both the // and the ! types.

Furthermore, it may be said from Fant's data that (2) the longer the length of the constriction, the higher the resonant frequencies, under the condition that the lengths of both the front cavity and the whole vocal tract are constnat. Thus, it may be said that the resonant frequencies of the ǂ type are higher compared with those of the / type. From the above, it may be qualitatively explained that the dominant frequency of click is higher for either / or ǂ compared with that of the corresponding click type in duration, i. e., // or !, respectively.

4-2. Frequency characteristics of the noise source

Besides the X-ray film, Traill measured the minimum pressure in the back cavity during the similtaneous period of both the primary and the secondary closures (1976, personal talk). Results were: -75 cmaq. for /, -65 cmaq. for ǂ and -70 cmaq. for //. The ! was not measured.

In his X-ray films taken during the above period, the capacity of the back cavity is gradually expanded from its minimum, and it reaches the maximum on the frame immediately before the primary release. Thus the minimum pressure value in the chamber must be reached immediately before the primary release, and the value is considered to be the pressure difference between the front and back cavity at the primary release. Under the assumption that the pressure value in the back cavity at its minimum volume is the same for each type of click, the minimum pressure value is in proportion to the degree of rarefaction, i.e., the ratio of the minimum volume to the maximum one. Furthermore, under the assumption that the cavity volume varies in proportion to the cavity area on the lateral X-ray film frame, the minimum pressure value in the back cavity may be estimated as about -50 to -60 cmaq. for the ! type. The values are shown in Table 4.

Click Noise Type / Area of Back Chamber	/	ǂ	//	!
Minimum area	23	20	58	54
Maximum area	300	98	399	160
Ratio of minimum to maximum (%)	7	20	15	34
Minimum pressure in back cavity (cm aq.)	- 75	- 65	- 70	-50 ∼ -60

Table 4 : The minimum pressure value in back chamber immediately before the primary release (after Traill, 1976, personal talk). The value of ! is estimated from the ratio of the minimum area of the cavity to the maximum one (see the text). Each capacity value of the area is relative to each other, and calculated from Traill's X-ray films.

Flanagan (1971) theoretically studied the production of turbulent noise, based on the experimental results by Heller and Windnall (1970) and by Gordon (1968, 1969). As to the frequency characteristics of the noise, he shows that the central frequency of the noise generated at the constriction is in proportion to V/D, where D means the "characteristic dimension" of the constriction and V the flow velocity through the constriction.

Under the assumption that the cross-sectional area of the constriction tion is the same for a ll types of click, the central frequency is in proportion to the pressure difference between the front and back cavity, since V is in proportion to the volume velocity U and U is in proportion to the pressure difference in certain regions of the pressure difference. Then, from measurements of the pressure value mentioned above, it may be considered that the central frequency of the click noise for the / or // type is higher compared with that of the ǂ or ! type, respectively.

4-3. Manner of primary release

On the basis of Beach's descriptions (1938), Chomsky and Halle (1968) tried to explain about the difference in duration of each click by their phonetic feature, "delayed primary release", which is defined acoustically.

Articulations causing these different acoustic effects were partly observed by Shibata (1968). He studied dynamicpalatographically the difference in articulation among Japanese [t], [ts] and [s] (uttered by a Japanese native speaker), and stated as follows: "In the initial phase of the articulatory movement (of [ts]), the manner of contact (between the tongue and the palate) is similar to that for [t]. The temporal pattern of release ... seems similar to that of [s]... " (Words in parentheses are the author's). That is, it is clearly observed that after the release, the period for both sides of the middle of the tongue contacting the palate is kept quite long for both [s] and [ts] compared with that for [t]. This seems to indicate that the height of the tongue tip mainly depends on that of the middle of the tongue for a short period of time after the release, and the release is an independent movement of the tongue tip. That is, in the cases of both [s] and [ts], the constriction after thr release exists longer, since the middle of the tongue is still kept in the high position, while in the case of [t], the constriction hardly exists because of the earlier downward movement of the middle of the tongue. The difference in manner of release mentioned above may apply to click types, though the difference in duration between the long click noise type (i. e., both / and //) and the short click noise type (i/e., both ǂ and !) is generally smaller than that between the affricates and the stops.

After the primary release, the ingressive airstream through the constriction may increase the air pressure in the back cavity, or in other words, it may decrease the pressure difference between the front and back cavities. If the decrease of the pressure difference exceeds some threshold value, the click moise may be ceased. Therefore, a bigger back cavity capacity may be required for the long click noise type than for the short click noise type. In fact, Traill's X-ray frame immediately before the primary release shows that the back cavity volume capacity of the / type is almost three times that of the ǂ type and that of the // almost two and a half times that of the ! (cf. Table 4).

During the period of the click, each strong intensity band (s) of both / and // type shifts down in frequency, though it is not observed in the case of either ǂ or ! , as mentioned earlier. This may be explained by a gradual increase in the cross-sectional area of the constriction and/or a decrease of the pressure difference between the front and back cavities.

Traill (1973) also measured the "length of release" of each type of clicks in the !xõ which belongs to the southern group of the San languages, "based on the recorded period of negative airlfow provided by Electro-kymographic tracing and those taken during two airflow studies [6]. " And he states as follows: "While these labels (concerning the two manners of primary release) many bear an intuitively obvious content and have been used regularly for non-click segments (e. g. [t] vs [ts]), instrumental investigation shows that terms which suggest difference in timing for the two kinds of release are quite inaccurate ... " (Words in the first parentheses are of the authro's.) In fact, the order of these lengths is not consistent in his measurements, but it can be seen that the length for / is longer compared with that for ǂ in both the electro-kymographic and airflow studies, and that for // is longer compared with that for ! in the electro-kymographic study. (In the airflow study, the length of ! is lacking.)

5. Concluding remarks

By the soundspectrographic analysis of clicks, it was found that the four types of clicks in Naron are acoustically distinguishable by both frequency characteristics and duration. The former may roughly distinguish both the / and ≠ types from the // and ! types, and the latter may distinguish the / and // types from both the ≠ and ! types.

The frequency characteristics may be roughly explained by the frequency characteristics of the noise-source and the vocal tract shape estimated from Beach's palatograms and Traill's X-ray films. The duration may depend on the manner of the primary release as mentioned earlier.

6. Acknowledgements

The author wishes to thank Prof. M. Sawashima, Director of the Research Institute of Logopedics and Phoniatrics, Univ. of Tokyo, for his helpful comments on this manuscript. The author also wishes to thank Dr. A. Traill, the Institute of African Studies, Univ. of the Witwatersrand, for his kindly teaching about the articulations of clicks from his valuable unpublished X-ray and pressure measurement data. All acoustic data were recorded in Botswana from January to June, 1976. The author is indebted to the Botswana Government for permitting him to conduct field-work, and the Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo Gaikokugo Daigaku (Tokyo Univ. of Foreign Studies) for their financial support.

Notes

(1) Ladefoged (1971) states that four principal airstream mechanisms are found in speech; the pulmonic egressive type, the glottalic egressive type, the glottalic ingressive type, and the velaric ingressive type, and that of the clicks is characterised as sounds caused by the velaric ingressive airstream mechanism.

(2) Beach (1938) details the articulation of each of four types of click noises (i.e., the /, ≠, // and ! types; cf. Table 1) of the Khoi (Hottentot).

(3) Some examples of the difference in the place of primary closure are shown in the following (except the bilabial type):

Researcher	language	/	≠	//	!
Beach (1938)	Hottentot	dental	dental	alveoalr	alveolar
Traill (1973)	!xõ (S.S.)	dental	dental	alveolar	alveo palatal
Maingard (1958)	!kɔ̃: (S.S.)	dental	alveolar		plato alveolar
Doke (1936)	≠khomani (S.S.)	dental	alveolar	(palato ?) alveolar	palato alveoalr
Tanaka (1978)	G//na & G/wi (C.S.)	dental	alveo-palatal	palatal	palatal
Bleek (1928)	Naron (C.S.)	dental or alveolar	alveolar		front palate
Snyman (1970)	!xũ (N.S.)	dental	alveolar	palatal	palatal

In the above, S.S. means that the language belongs to the southern group of the San languages, C.S. to the central group and N.S. to the northern group. It can be seen in the above that these places of primary closure are roughly divided into two groups, i.e., one consisting of both / and ≠ and the other of both // and ! .

(4) The difference in manner of primary release is expressed as follows: e.g., affricate vs implosive (Beach, 1938), fricative vs occlusive (Trubetzkoy, 1969), delayed release vs spontaneous release (Chomsky and Halle, 1968), or with friction vs without friction (Traill, 1973).

(5) In Naron, Bleek (1928) found eleven phonetic accompaniments for each type of click. But, these may be grouped into six distinctive accompaniments, i.e., voiced, voiceless (plain), aspirated, glottalized, ejected, and nasalized accompaniments (Kayga, 1978).

(6) In his studies, the number of samples is about 13 for each type of click. Unfortunately, characteristics of his instrumental system are not detailed there.

References

Beach, D.M. 1938. The Phonetics of the Hottentot Language. Cambridge: W. Heffer & Sons, Ltd.

Bleek, D.F. 1928. The Naron. Cambridge Univ. Press.

Chomsky, N. and M. Halle, 1968. The Sound Pattern of English. New York: Harper & Row.

Doke, C.M. 1936. "An Outline of ≠khomani Bushman Phonetics", Bantu Studies, Vol.10, No.4, pp. 433-460.

Fant, G. 1960. Acoustic Theory of Speech Production. Mouton & Co.

Flanagan, K. N. 1971. "Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations, " The Journal of the Acoustical Society of America, Vol. 50, No.4 (Part 2), pp.1180-1192.

Gordon, C.G. 1968. "Spoiler-Generated Flow Noise I. The Experiment". The Journal of the Acoustical Society of America, Vol.43, pp. 1041-1048.

——————. 1969. "Spoiler-Generated Flow Noiee II. Result." The Journal of the Acoustical Society of America, Vol.45, pp.214-123.

Heller, H. H. and S. E. Widnall, 1970. "Sound Radiation from Rigid Flow Spoilers Correlated with Fluctuating Forces." The Journal of the Acoustical Society of America, Vol.47, pp. 924-936.

Kagaya, R. 1978. "A Phonetic Sketch of Naron around Ghanzi." The Journal of Asian and African Studies, The Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo Gaikokugo Daigaku, No. 16. (in press).

Ladefoged, P. 1971. Preliminaries to Linguistic Phonetics. Chicago: The Univ. of Chicago Press.

Lanham, L. W. 1964. "The Proliferation and Extension of Bantu Phonemic Systems Influenced by Bushman and Hottentot." In the Proceedings of the 9th International Congress of Linguistics, edited by H. Lunt, pp. 382-391. Mouton & Co.

Maingard, L. 1958. "Three Bushman Languages, Part II: The Third Bushman Language." African Studies, Vol. 17, No. 2, pp. 100-115.

Shibata, S. 1968. "A Study of Dynamic Palatography." Annual Bulletin, No. 2, pp. 28-36. Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, Univ. of Tokyo.

Snyman, J. W. 1970. An Introduction to the !xũ Language. Cape Town: A. A. Balkema.

Tanaka, J. 1978. A San Vocabulary of the Central Kalahari. ILCAA African Languages and Ethnography 7. The Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo Gaikokugo Daigaku.

Traill, A. 1973. "A Preliminary Sketch of !xõ Bushman Phonetics." Work in Progess, Linguistic Dept. of Edinburgh Univ., No. 6.

Trubetzkoy, N. S. 1969. Principles of Phonology. translated by C. Baltaxe. Los Angeles: Univ. of California Press.