CHOKRI (PHEK DIALECT):
PHONETICS AND PHONOLOGY

Brian Bielenberg and Zhalie Nienu
University of California, Berkeley

INTRODUCTION

This paper examines the phonetics, phonemics and phonology of the Phek dialect of the Chokri language as spoken by a particular individual now living in the United States. The main purpose is to develop a more complete understanding of the phonology of the language than has previously existed. Throughout the paper comparison will be made with data presented by Marrison (1967) and Nienu (1990), the only known sources on Chokri phonology. Attention will also be paid to the data on Angami (Kohima) presented in Marrison as a means of estimating the degree of linguistic interference resulting from the consultant's complicated linguistic history, which will be outlined below.

The Language

Chokri is one of seventeen Naga languages spoken in the northeastern Indian state of Nagaland (SIL 2000). The Naga languages are members of the Kamarupan branch of the Tibeto-Burman phylum of the Sino-Tibetan language family. Chokri is also known as Eastern Angami, Chakrima Naga, Chakrü, and Chakhesang.[1] In 1991, the most recent data available, it was estimated that there were approximately 20,000 native speakers of Chokri, with Cheswezumi (Chazouba) being the main village where Chokri is spoken (SIL Website 2000). Chokri is considered by many to be a dialect of Angami, the dominant language of western Nagaland with over 100,000 speakers. The Chokri are categorized as an eastern Angami group, but there exists a large population of Chokri speakers in Kohima, the largest city of western Nagaland. Cheswezumi lies approximately 40 miles southeast of Kohima, and Sohima, the informant's village of birth, lies another 50 miles further southeast. It is worth noting that Marrison (1967) collected his Chokri data from Cheswezumi. The consultant readily recognized this after examining the forms listed in Marrison. He noted that his uncle, from Sohima, is married to a woman from Cheswezumi, and that while their dialects

---

[1] Chakhesang is actually a sort of *lingua franca* incorporating elements of Chokri (chak-), as well as of the closely related Khezha (-khe-) and Sangtam (-sang). [Ed.]

are mutually intelligible, there do exist a number of lexical differences. For instance, the Phek dialect of Chokri uses the word /tsɨ/ 'small', whereas Cheswezumi dialect speakers use /kəɲɪ/. Interestingly, although Sohima is further away geographically than Cheswezumi from the center of the Kohima dialect of Angami, its dialect appears to have more cognates within this consultant's speech. This may be a result of time spent in Kohima by the consultant.

The Consultant

The speaker consulted for this investigation has a complicated linguistic history, but all evidence points to his being a fluent speaker of the Phek dialect of Chokri. He was born and raised in the Chokri speaking village of Sohima, a small village of approximately 100 houses. The consultant reports that most people living in Sohima are bilingual, speaking both Chokri and Bochiri. Many residents speak Angami and English as well, as they are the languages of schooling and the church. For most of the informant's life Chokri was the main language spoken in the home and with his peers, even though he attended English medium school in the capital city of Kohima. In Kohima, a predominantly Angami speaking city, the consultant lived with his uncle (from Phek) and grandmother, both speakers of the Phek dialect of Chokri, and this was the language of the home. He also intermittently spent time in Phek living with another Chokri speaking paternal uncle.

During his time in Kohima, the consultant recalls that he spent many of his hours outside of school in a Chokri speaking part of town interacting with Chokri speaking people. However, he also spent enough time interacting with Angami (Kohima) speakers in order to learn that language as well. Approximately eight years ago, at the age of 24, the consultant came to the United States to further his education. He has recently completed a Masters of Divinity degree at a California seminary and is now working as a youth leader for a church in California. His father also lives in California and they speak Chokri on a regular basis. Although at first the consultant's knowledge of Chokri appeared a bit "rusty," it was only a short time before he was constructing complex sentences and relating narratives. Still, after 9 months of serving as a consultant, he continues to have some difficulty recalling "basic" words such as *boat, riverbank, spoon, eggplant*, and *drum* when elicited out of context. The consultant appears to have a good deal of meta-knowledge of languages and was able to point out differences between Chokri and Angami, as well as between the northern and southern dialects of Chokri. He notes that Angami (Kohima) and Chokri (Phek) have many similarities phonologically and lexically, but that they are not mutually intelligible. Comparison with Marrison's data confirms that there is a high percentage of

cognacy in basic vocabulary. The informant's speech is clear and carefully pronounced, with no obvious speech defects to affect his pronunciation.

DATA COLLECTION

The data for this paper were collected over a period of 9 months in two two-hour sessions each week as part of a linguistics course on field methods. Elicitation began with the names of natural objects, numerals, colors and kinship terms. It progressed to simple phrases and sentences with focus on particular aspects of grammar. By the middle of the second semester narratives were being elicited, with a total of eight texts generated. Transcription and translation of two of the more extensive stories are included in Appendix A. The first author also met individually with the consultant, the second author, outside of class on a weekly basis to elicit further information, especially with regard to tone. Although Marrison (1967) contains over 500 Chokri glosses with forms, none of them have tonal markings and this is considered a great weakness of his data. A large part of the current work is concerned with obtaining accurate tonal markings for the words elicited. To this end, all new words encountered in class were re-elicited outside of class and their tonal structure identified. Additionally, lists were regularly generated in order to allow for similar syllable structures, either by coda, which in this language consists only of a vowel, or by tone, to be cross-checked against one another. The total number of forms elicited and used for phonemic analysis in this paper is 621. After each session the data were entered into a database so as to simplify future sorting tasks. The data were entered in such a way as to allow for sorting by gloss, form, number of syllables, tone(s), first consonant, first vowel, second consonant, second vowel, and so on. Columns were also included for the Chokri (Cheswezumi) and Angami (Kohima) forms from Marrison's data to allow for easier cross-referencing.

The majority of forms elicited, approximately 95%, consisted of one or two syllables. The remaining 5% of forms were trisyllabic. Such a lexical structure is expected for a member of the Sino-Tibetan language family. A few quadrisyllabic words were also identified, but are not used in this analysis, as they tended to be compound words and showed no phonemic variation from the uncompounded roots. The canonical structure of syllables is $(C)(r)V(Tone)$, where the initial consonant and tone are optional. In a number of words syllabic **n**'s, **m**'s and **r**'s were also encountered, but in every case these are the result of a reduction process. The consonantal clusters observed were [kʵ kʰʵ pʵ pʵʰ], as seen in the following four examples:

[kɹ̀í]   'to fall'
[kʰɹ̀í]   'moon'
[məpɹ̂ê]   'be scared'
[dzɨpɹ̂ʰɔ̃]   'break something by touching'

Several affricates were also encountered and will be discussed in their own section. For monosyllabic words, the syllable always appeared tone-bearing, even in natural speech. A monosyllable rarely, if ever, reduces down to an unstressed vowel. It is difficult to know whether Marrison reached the same conclusion, as he appears to use /ü/ for both stressed and unstressed central vowels. The case of polysyllabic words is different, with the vowel of the non-final syllables often being reduced and non-tone bearing. Many of the first syllables in such words are bound or unbound prefixes that perform various functions. In almost all cases the prefixes are non-tone bearing or have a 22 or 33 (Chao pitch scale) tone, perhaps a neutralization of the two falling mid-tones (see below).

After identifying the phonemes of the language, the data were examined for gaps. To determine whether these gaps were accidental or actual, exploratory elicitation was used. In repeating words to the consultant for verification, I often intentionally mispronounced in order to determine the range of possible pronunciation.

ANALYSIS

Analysis of Chokri involves not only a detailed examination of the segmental phonemes of the language, but also of the tones. Two words differing by a single phoneme can only be considered to be a minimal pair if they have identical tone structures. Thus it was important early on to devise a simple method of marking tone.

Four tones were identified and they are marked using the following symbols: ` for a low tone, ~ for a low-mid tone, ^ for a high-mid tone, and ´ for a high tone. An allotone of the high-mid tone was also found (see below). The equivalents of these tonal markings in terms of Y.R. Chao's tone letters are shown in Table 1. The typology of this tonal system is unusual in that it contains four falling tones, with the two mid-tones overlapping in an interesting manner, the difference between them being that the high-mid tones drop more rapidly during the second half of pitch duration than the low-mid tones. (See below for more details). Syllables without a full vowel never appeared to be tone bearing, so it was decided that these would be left unmarked. All the symbols used for the consonants and vowels are those of the International Phonetic Alphabet. Where