NON-NATIVE SPEAKERS' IDENTIFICATION OF LEXICAL TONE CONTRASTS

Kofi Adu Manyah

Kwame Nkrumah University of Science and Technology, Kumasi

ABSTRACT

This study examined how native speakers of French, a non-tone language with different lexical stress properties, identified two Twi tones in a series of monosyllabic lexically contrasting words. These words were produced, first in isolation and second, in a carrier sentence. Three experiments were conducted. Experiment 1 tested identification of tone contrasts by native speakers of Twi. Experiment 2 tested native French speakers' identification of tones presented in isolated context and Experiment 3 tested identification of tones presented in a carrier sentence. Acoustic investigations reveal that pitch difference is the determining factor in distinguishing high and low tone contrasts in Twi. Results also show that the identification rate of tones in an isolation context was higher than the identification rate of tones in a carrier sentence context. Findings imply that linguistic experience with non-native tones does not necessarily preclude tone identification.

Keywords: lexical tones, non-native speech identification.

1. INTRODUCTION

Twi is a register tone language with a two-tone system (high/low) and the downstep. The language uses tone for lexical and grammatical distinctions. There is a contrast between 2 pitch heights, and this occurs in words where the assigned syllable pitch is relatively higher or lower. All vowels (oral and nasal) and syllabic consonants carry tones. It has verbal tones (a class of low tone verbs and a class of high tone verbs), nominal tones, adjectival tones, adverbial tones and homotones. In the category of homotones are words that exhibit the same tone patterns but have different meanings [5, 6]. Where there is verb serialisation involving 2 different stems (i.e. verb+verb constructions) the 1st stem carries a low tone [̀] and 2nd stem has a high tone [́]. In verb reduplication, the 2 stems carry low tones [1, 2].

There are two possible tone patterns in monosyllabic words: High and Low. Register tone patterns in disyllabic morphemes are High-High, High-Low, Low-High and Low-Low. Register tone patterns in trisyllabic morphemes are: Low-Low-Low, Low-Low-High, Low-High-High and High-High-High.

In terms of production, stress in French is consistently marked by an increase in duration, whereas F0 plays an important role in terms of perception [7]. Empirical research has shown that French listeners do rely on certain acoustic cues in the rhythm of the language to segment speech [12]. How effective can these apply to the perception/ identification of Twi lexical tones?

Aim

This study reports observations and results obtained in an experiment on the identification of two lexical tone contrasts by non-native Tone Language Speakers viz. French speakers.

1. F0 values of the target vowels in mono syllabic morphemes were obtained in order to verify that tone contrasts indeed rely on factors pertaining to pitch.
2. F0 values were calculated for the two sets of tone assignments.
3. Intensity values were calculated for the two sets of tone assignments.
4. Durational values were taken for the two sets of tone assignments.
5. Statistical analyses (ANOVAs) were carried out on all measures obtained from the speaker ($p \leq 0.01$).

2. METHOD

Two adult male native speakers of Twi, with no speech or hearing impairment, served as speakers. The data consisted of acoustic productions of a series of minimal pairs, containing high and low tones in (a) isolation and (b) a carrier sentence, at a self-selected conversational rate. The randomised list of the stimuli in isolation was produced at least 5 times by each speaker

Material: The stimuli in isolation consisted of monosyllabic lexically contrasting words differing in pitch. Tonal distinctions in Twi are not category limited so stimuli were not considered from the same category i.e. (n) vs. (n), (v) vs. (v), (adj vs. adj), etc.

1. **da** sleep (v), day (n) /dá/ vs. /dà/ never (adv)
2. **dua** tree (n) /dʊá/ vs. /dʊà/ plant (v)

3. **bra** life, existence (n) /brɑ́/ vs. /brɑ̀/come (v)
4. **srɛ** thigh (n) /srɛ́/ vs. /srɛ̀/ beg (v).

All the stimuli, embedded in a carrier sentence, **kã sɛ ..... kyerɛ me** [Say……to me], were also produced 5 times by each speaker:

kã sɛ /**dɑ́**/ kyerɛ me *vs.* kã sɛ /**dɑ̀**/ kyerɛ me
kã sɛ /**dɥɑ́**/ kyerɛ me *vs.* kã sɛ /**dɥɑ̀**/ kyerɛ me
kã sɛ /**brɑ́**/ kyerɛ me *vs.* kã sɛ /**brɑ̀**/ kyerɛ me
kã sɛ /**srɛ́**/ kyerɛ me *vs.* kã sɛ /**srɛ̀**/ kyerɛ me.

Acoustic data were recorded in an anechoic room for one speaker after the preliminary training session. By means of PRAAT sound editor, Fundamental frequency (F0) and Intensity measurements were taken for the tones in isolation and for the tones in a carrier sentence. Durational measurements of tones were also done for the:

a) target vowel
b) syllable CV

F0, Intensity and Duration are key because it is widely known that a prominent syllable is marked by variations in four acoustic cues: fundamental frequency (pitch), amplitude (volume), duration (length), and formant structure (a different timbre of the vowel). Of the four cues F0 seems the most important in terms of perception in French [7].

In the Twi language, like many tone languages in E. Asia and Africa, there is also a one-to-one association between tone and syllable i.e. each monosyllabic word or morpheme is associated with a tone [13]. A high tone is said to be distinguished not only by acoustic features but also has to be considered as a perceptual phenomenon. Therefore, in defining it one must account not only for its production but also for its perception by the interlocutor [7]. How do non-tone language listeners perceive the unfamiliar tonal distinctions of a lexical tone language? Given that the amount of linguistic experience in tonality use does not guarantee a facilitating effect for the native listeners of a different tone language, viz. Cantonese listeners of Mandarin tones [10]. The core question of the present study, therefore, was how Twi lexical tones may be identified by listeners of a non-tone language, French, in terms of isolated words and at the sentence level.

Experimental Design

Three experiments were performed to find an answer to the question. Experiment 1 tested identification of lexical contrasts based on tones by native speakers of Twi. Experiment 2 tested the identification of lexical tone contrasts by native French speakers presented in isolated context. Experiment 3 tested the French speakers' identification of lexical tones presented in a carrier sentence.

Experiment 1: Identification Test

This experiment sought to test native speaker lexical tone identification of stimuli produced in isolation and in a carrier sentence. Three Ghanaian female students (ages 21, 22, 27: average age 23.3) were recruited to accomplish the task.

Procedure: A subject listened to the test sound (3 times) with a PRAAT sound editor through *Sennheiser* headphones, and decided, which of the two meanings (choices) printed on an answer sheet corresponded to the 1st word and which corresponded to the 2nd word. They marked the choice on the answer sheet accordingly. A five (5) second interval was given after the stimuli to enable the subjects to indicate their level of confidence (on a scale ranging from 1-5). A short training session consisting of 2 examples each of discrimination tests in each category (isolation & carrier sentence) preceded the test.

Results: The accuracy of lexical tone identification of stimuli produced in isolation and in a carrier sentence was displayed in this experiment. The high (lexical) tone was heard as high tone 100% of the time in the isolated context and 100% in the carrier sentence context. The low (lexical) tone was heard as a low tone 100% of the time in the isolated context and 100% of the time in the carrier context. With data provided by acoustic investigations, viz. F0 on pitch difference, I could conveniently subject the corpus to the listener identification process and interrogation on high and low tone contrasts in Experiments 2 and 3.

Experiment 2: The experiment tested French native speakers' identification of tones in a carrier sentence. A total of thirty-one (31) third year students (ages between 19 and 30 years: average age 21.3), at the time of the study, participated in the lexical tone identification test. To best enable us obtain answers to the questions, Purposive Sampling was used in selecting the subjects. The students had no previous exposure to the Twi Language and its tonal structure.

Procedure: The stimuli were played in random order, each repeating three times in immediate succession. Each pair of stimuli was played with a bip sound preceding each sequence. An answer sheet was prepared on which, corresponding to each pair, there were 2 boxes representing a high and a low tone.

The subjects' task was to identify the tone of each target, marking the appropriate tone on the answer sheet. They were given 5 seconds after each

repetition to indicate on the same sheet their level of confidence on a scale, ranging from 1 (less confident) to 5 (very confident). The process was repeated four times in random order.

Experiment 3
This experiment was set to test speaker identification of tones in a carrier sentence. The same number of French native speakers, with no previous exposure to Twi tones, took part in the experiment.
Procedure: The same stimuli was used but were inserted in a carrier sentence.

Results: Experiment 2 and Experiment 3
Listeners' response patterns, or performance rates of stimuli discrimination, differ depending on whether a target tone was produced in isolation or in a carrier sentence context. When the context was in isolation, the majority of the tokens were identified correctly as having high or low tone (56.5%). When the context was in a carrier sentence, the tokens were identified correctly as having high or low tone (48.8%). They have relatively low standard deviations of 2.3% and 6.7% respectively. The identification rate of the tones in an isolation context was therefore higher than that of the tones in a carrier context (Table 3). These results are in agreement with similar results attested for Mandarin [14].

3. RESULTS AND DISCUSSION

Table 1: Average F0, Intensity and Durational values with standard deviations for stimuli produced in isolation

Stimuli		L F0 (Hz)	L Int (dβ)	L Dur (ms)	H F0 (Hz)	H Int (dβ)	H Dur (ms)
da	Av	135	80	224	172	87	269
	Sd	7	1	39	17	0.5	10
bra	Av	134	77	261	180	82	309
	Sd	6	3	28	4	4	20
dua	Av	134	78	322	174	81	353
	Sd	6	4	27	8	4	31
srɛ	Av	163	72	340	183	77	406
	Sd	6	2	14	12	2	17

For Experiment 1, the basic trend in the data, F0, intensity and durational values and standard deviations for the two sets of tone assignments are presented in Table 1 and Table 2. Data were obtained by averaging over the individual values of those measurements for all five tokens of the word. The overall data indicate that the most important parameter for determining tone assignment contrasts, i.e. relative pitch, is highly significant ($p<0.001$).

Table 2: Average F0, Intensity and Durational values with standard deviations for stimuli produced in a carrier sentence

Stimuli		L F0 (Hz)	L Int (dβ)	L Dur (ms)	H F0 (Hz)	H Int (dβ)	H Dur (ms)
da	Av	135	71	236	157	76	254
	Sd	19	2	26	23	2	41
bra	Av	118	68	239	139	74	329
	Sd	12	2	7	13	2	23
dua	Av	127	71	305	150	75	376
	Sd	3	1	22	13	2	5
srɛ	Av	119	67	281	183	72	327
	Sd	9	1	20	9	2	36

Table 3 below illustrates the general tendency of the results i.e. the success rate of identification of the two lexical tones by the individual speakers, in isolated word and carrier sentence contexts. It also shows their level of confidence and the standard deviations for each contrast.

The majority of the tokens were identified correctly as having high or low tone (56.5%) when presented in isolation (low speaking rate) but as having high or low tone (48.8%) in the carrier sentence context (high speaking rate). The results seem to lend credence to the claim that whether a particular tone in a language is inherently dynamic or composed of static elements can be seen more distinctly only when speaking rate is slowed down so that there is sufficient time for the underlying targets(s) to be fully implemented [13]. In other words, in this study, identification rate of tones in an isolation context was higher than the identification rate of tones in a carrier sentence because speaking rate is slowed down in an isolation context.

According to the average figures, the level of confidence displayed by the listeners in Table 3 is higher in the isolated context than the carrier sentence context. It can also be observed that the difference between the level of confidence indicated by the listeners and the final result obtained is practically identical in the two contexts. The difference is 33.8% in the isolated word context and 32.8% in the carrier sentence context.

However, intra syllabic comparison of stimuli reveals that listeners' performance rate is not the same for the stimuli in isolation and the stimuli in a carrier sentence. In fact, the performance rate is lower for **bra** and **srɛ** compared to **da** and **dua** when they appear in a carrier sentence.

On the other hand, when tokens are presented to the listeners in isolation the performance rate is higher for **da** and **bra** compared to that of **dua** and **srɛ.**

The general tendency is that the syllabic durational values of tones produced in a carrier sentence are less than the syllabic durational values produced in isolation. This could be due to the difference in speaking rate between the two conditions i.e. a slower speaking rate for stimuli produced in isolation (Tables 1 & 2). There is a high percentage (88%) of data on absolute durational values confirming the hypothesis that tones produced in a carrier sentence would be shorter in duration than those produced in isolation indicating a difference in speaking rate.

Table 3: This table shows the subjects' response patterns of identification: lexical tone contrasts in isolation and in carrier sentence contexts

	Age	Isolated Word %	Level of Confid %	Carrier Sentence %	Level of Confid %
	30	61.9	83.6	47.6	66.4
	20	52.4	96.2	47.6	83.6
	20	57.1	93.3	57.1	95.2
	21	57.1	83.8	57.1	57.1
	21	57.1	75.2	57.1	56.2
	21	57.1	89.5	47.6	75.2
	23	57.1	100	38.1	97.1
	20	52.4	100	52.4	100
	21	57.1	96.2	47.6	78.1
	21	57.1	96.2	38.1	80.1
	21	57.1	95.2	57.1	91
	22	57.1	83.8	57.1	96.2
	21	52.4	94.3	47.6	98.1
	20	57.1	89.5	38.1	92.4
	23	57.1	94.3	52.4	83.6
	19	57.1	89.5	47.6	79
	20	57.1	83.8	38.1	76.2
	20	57.1	94.2	52.4	82.9
	20	57.1	96.2	52.4	84.4
	22	52.4	90.9	38.1	83.6
	19	57.1	89.5	52.4	91
	20	57.1	75.2	47.6	57.1
	23	61.9	100	38.1	97.1
	22	57.1	94.2	47.6	82.9
	22	52.4	94.3	38.1	83.6
	21	57.1	83.8	57.1	75.2
	24	57.1	96.2	52.4	83.6
	20	57.1	75.2	47.6	57.1
	21	52.4	83.8	57.1	79
	23	57.1	95.2	52.4	83.6
	20	57.1	83.8	47.6	84.4
Av	21.3	56.5	90.2	48.8	81.6
StDv	2	2.3	7.1	6.7	12.2

4. CONCLUSIONS

This investigation has shown, on the basis of the selected corpus and evidence from our acoustic data that pitch difference (F0) is indeed one of the key determining factors in distinguishing high and low tone contrasts in Twi.

The identification test undertaken to verify pitch contrast for the two phonological classes, reveals that tone differences really determine lexical contrasts.

French subjects 'succeed' in identifying Twi lexical contrasts based on tones when tones are produced in isolation at a performance rate of 57% and when tones are produced in carrier sentences at a rate of 49%, for monosyllabic words. This could be explained by the fact that speaking rate is slowed down (in isolation) and there is sufficient time for elements to be seen more distinctly.

Intra syllabic phonemic comparison, suggests that performance rate is not the same for stimuli produced in isolation and stimuli produced in a carrier sentence.

Durational values of tones produced in isolation are generally greater than those produced in a carrier sentence due to the difference in speaking rate between the two conditions.

Further investigations will focus on identification performances for polysyllabic (disyllabic, etc.) tone contrasts with data based on more native French listeners.

5. REFERENCES

[1] Adu Manyah K. 2010. *Parlons Twi*. Langue et culture. Paris: l'Harmattan.

[2] Adu Manyah K. 2007. Considérations tonologiques du twi. *XXVIIèmes Journées d'Étude sur la Parole de l'AFCP*, 273-276.

[3] Best C. T., Bohn P. O, Faber A. 2003. Cross-Language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages, *15th ICPhS*. Barcelona. 2889-2892.

[4] Chang S-E. 2008. Perception of South Kyungsang Korean tones. *Proceedings of Speech Prosody*, 277-280.

[5] Dolphyne F. A. 1988. *The Akan (Twi-Fante) Language. Its Sound Systems and Tonal Structure*. Accra, Ghana Universities Press.

[6] Dolphyne F. A. 1986. Tone and Grammar in Akan; the tone of possessive constructions in the Asante dialect. Phonological Representation of suprasegmentals, *Publications in African Lang. and Linguistics.* 4: 35-49.

[7] Frost, D. 2011. Stress and cues to relative prominence in English and French: A perceptual study. *J. of the Int. Phonetic Assoc*. Vol. 41, No. 1, 67–84.

[8] Man Gao. 2016. Perception of Lexical Tones by Swedish Learners of Mandarin. *Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition at SLTC*. 34-40.

[9] So, C. K. & Best, C. T. 2014. Phonetic influences on English and French listeners' assimilation of mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36, 195-221.

[10] So, C. K. 2006. Perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Proc. of the 11th Australian Int. Conf. on Speech Sc. & Tech*, ed. Paul Warren & Catherine I. Watson.

[11] Tran D. D., Castelli E., Serignat J-F, Trinh V. L, Le X. H. 2005. Influence of F0 on Vietnamese syllable perception. *Interspeech 2005-Eurospeech*. 1697-1700.

[12] Wioland, F. 1991. *Prononcer les mots du français. Des sons et des rythmes*. Paris: Hachette.

[13] Xu, Y. 2004. Understanding tone from the perspective of Production and Perception. *Language and Linguistics*. 5.4: 757-797.

[14] Y. 1994. Production and perception of coarticulated tones, *J. Acoustical Society Of America*. 2240-2253.