

# ACOUSTIC AND PHONOLOGICAL FACTORS IN THE PERCEPTION OF ENGLISH /r/ AND /l/ BY JAPANESE LISTENERS

John S. Logan, Kathy BharrathSingh  
Department of Psychology, Carleton University, Ottawa

## 1. INTRODUCTION

Japanese listeners generally have difficulty perceiving the difference between /r/ and /l/, even after years of exposure to English (MacKain, Best, & Strange, 1981). Such a finding suggests that substantial reorganization of the perceptual system is required in order for Japanese listeners to achieve fluent perception of /r/ and /l/. However, Logan, Lively, and Pisoni (1991) and Lively, Logan, and Pisoni (1993) demonstrated that Japanese listeners can learn to identify /r/ and /l/ more accurately after completion of a laboratory training procedure. Their procedure utilized a large ensemble of stimuli consisting of words produced by five talkers that contained /r/ and /l/ in five phonetic environments. Performance gains were modest but reliable after three weeks of training. Although this work indicated that adult Japanese listeners were capable of learning about /r/ and /l/, additional questions remained concerning how training affected the structure of phonemic categories.

The work of Logan et al. (1991) and others indicated that listeners can learn to redirect their attention to previously unattended portions of the speech signal. What factors affect the allocation of attention? One factor that seems to play an important role when Japanese listeners learn to perceive /r/ and /l/ is the phonetic environment in which /r/ and /l/ are embedded: performance is better when the phonemes are located in word-final positions than in word-initial positions. As a result of acoustic analysis, several researchers have proposed that this effect is due to differences in the duration of /r/ and /l/ in these two contexts. If the phonemes are located in word-final positions, they tend to be longer in duration relative to the same phonemes located in word-initial positions (Dissosway-Huff, Port, & Pisoni, 1982; Sheldon & Strange, 1982). As a consequence, listeners may have more opportunity to extract the critical information denoting /r/ or /l/ in word-final positions than in word-initial positions.

However, it is important to note that duration is not the only factor that may play a role in the ease with which listeners can direct their attention. Henly and Sheldon (1986) found that native speakers of Cantonese differed from Japanese listeners regarding the effect of phonological context on identifying /r/ and /l/. Whereas Japanese listeners had difficulty identifying /r/ and /l/ in initial positions, Cantonese listeners showed difficulty identifying /r/ and /l/ in final positions, indicating the critical role of the listener's phonological system. The difference between Japanese and Cantonese listeners suggests that one way to study how listeners redirect their attention is to experimentally manipulate the duration of the speech signal that corresponds to /r/ or /l/ and present the stimuli to Japanese and Cantonese listeners. By directly comparing the performance of these two groups, it should be possible to determine the relative contribution of external acoustic cues and internal phonologically-determined attentional mechanisms.

As a first effort towards directly investigating the effect of /r-/l/ duration on perception, a group of English and Japanese listeners were presented a series of English words. Stimuli with word-initial /r-/l/ had the /r/ or /l/ portion elongated, whereas stimuli with word-

final /r-/l/ had the /r/ or /l/ portion shortened. If duration of the /r-/l/ segments is the primary determinant of performance, then the Japanese listeners should identify the elongated stimuli in word-initial position with greater accuracy than normal duration word-initial stimuli. Conversely, they should find the shortened versions of the stimuli containing word-final /r-/l/ more difficult to identify than the normal versions. If their performance is not positively affected by the duration manipulation, it would suggest that the phonological/attention system is so firmly entrenched that simple efforts to make the acoustic features underlying /r/ and /l/ more or less salient are unlikely to succeed.

## 2. METHOD

Stimulus preparation and experimental control utilized CSRE software (Jamieson, Ramji, Neary, & Baxter, 1990). A male speaker of Canadian English (CE) recorded tokens of the following /r-/l/ minimal pairs: late-rate, lead-read, lock-rock, loot-root, fear-fill, hair-hell, more-mole, and tall-tar. Tokens were low-pass filtered at 4.8 kHz and then digitized at 10 kHz using a 12-bit A/D converter. Tokens were then modified in one of two ways. Tokens containing word-initial /r-/l/ had the duration of the /r/ or /l/ portion of the signal elongated, whereas tokens containing word-final /r-/l/ had the duration of the /r/ or /l/ portion of the signal shortened. In the case of tokens containing word-initial /r-/l/, the longest steady-state portion within the /r-/l/ portion of each token was determined using a waveform editor. This section was iteratively reduplicated to produce two versions of each token in addition to the original, one that was 40 ms longer, the other that was 80 ms longer. In the case of words containing word-final /r-/l/, 20 ms and 40 ms sections of the steady-state portion within the r-l portion of each token were removed. Altogether, 48 unique stimuli were produced using this procedure.

The stimuli were first tested with a group of six native speakers of CE, all of whom reported no history of a speech or hearing disorder. The pretest consisted of a two-alternative forced choice (2AFC) identification task in which each stimulus was presented once. Stimuli were presented over a Radio Shack loudspeaker at approximately 75 dB SPL. Listeners were seated in front of a computer monitor that displayed two words from a minimal pair. 250 ms later the auditory stimulus was then presented. Subjects were then required to move a mouse to indicate which word from the minimal pair had been presented. As well, each CE listener rated the quality of the 48 tokens using a 7-point scale, where 1 was "very clear", 4 "clear", and 7 "distorted".

Two Japanese subjects were recruited from the Carleton University community. One was a graduate student, the other an undergraduate. Both participants were first exposed to English in a formal setting at age 13 when they began English classes in school. Each had lived in Canada for approximately five years at the time they were tested. Both reported that although English was their primary language for educational purposes, they used Japanese in their non-student activities. Both listeners reported no history of a speech or hearing disorder.

The procedure used to test Japanese subjects was essentially the same used to test the CE listeners. The only changes were the

use of headphones (Sony MDR-P1) to present stimuli, an increased number of stimulus repetitions (8 x 48), and no requirement to rate the stimuli. In addition, the Japanese listeners were asked to complete a language experience questionnaire.

Table 1 – Mean quality ratings of stimuli by CE listeners as a function of stimulus modification (original, moderate, or extreme), position (initial and final), and identity of r-l segment (/r/ versus /l/).

	Original	Moderate	Extreme	Mean
/r/-Final	2.3	2.1	2.5	2.3
/r/-Initial	1.7	2.3	3.3	2.4
/l/-Final	2.0	2.2	2.0	2.1
/l/-Initial	1.9	2.1	2.5	2.2
Mean	2.0	2.2	2.6	2.3

Table 2 – Mean identification performance for Japanese listeners as a function of stimulus modification (original, moderate, or extreme), position (initial and final), and identity of r-l segment (/r/ versus /l/).

	Original	Moderate	Extreme	Mean
/r/-Final	98.4	75.0	73.4	82.3
/r/-Initial	75.0	90.6	78.1	81.3
/l/-Final	92.2	87.5	90.6	90.1
/l/-Initial	57.8	48.4	42.2	49.5
Mean	80.9	75.4	71.1	75.8

### 3. RESULTS

Overall, each token was accurately perceived by the CE listeners, with no subject making an identification error. However, ratings appeared to be more sensitive to differences among the stimuli than the accuracy data. Table 1 shows the mean ratings across duration (original, moderate, & extreme), /r-/l/ position (initial & final), and segment identity (/r/ or /l/). Tokens were generally perceived as clear, with a mean rating of 2.3. An analysis of variance revealed no reliable effect of /r-/l/ position (initial vs. final) nor any effect of whether the word contained /r/ or /l/ ( $p > 0.05$ ). However, a significant main effect of duration,  $p < 0.05$ ,  $F(3,10) = 4.695$ ,  $MSE = 2.02$ , was obtained. As the duration of the change increased (by either the addition or deletion of a portion of the stimulus), subjects rated the modified stimuli slightly more distorted. The effect was small, but it was observed across all six listeners. No other reliable effects were obtained.

Mean identification accuracy for the Japanese listeners is shown in Table 2. Overall, their performance was substantially worse than the CE listeners, with 75% of the stimuli correctly identified by the Japanese listeners compared to 100% correctly identified by the CE listeners. Because of the small number of Japanese listeners, it was not possible to carry out any statistical tests on the data. However, an examination of Table 2 indicates several interesting findings. First, identification performance is poorer for /r-/l/ in word-initial environments (63%) than for word-final environments (86%), replicating earlier work. Second, performance for stimuli containing word-initial /r-/l/ is substantially worse than in any other environment (50% compared to between 81% and 90%). Finally, the results indicated that subjects' performance did vary as a function of duration, although not always in the direction predicted. Only for /r/-initial words did lengthening the signal appear to have any positive effect, and that was limited to the 40 ms increase. For the length decrease in word-final position, the only noticeable effect appears in /r/ final words. In the other two environments the duration modification appeared

to have either little effect (/l/-final) or an effect opposite to that predicted (/l/-initial).

A comparison of the CE and Japanese listeners requires rating data to be compared to accuracy data. Though indirect, the two tasks should each be measuring some aspect of how listeners perceive the stimuli. Overall, the two sets of data were not correlated ( $r = -0.026$ ). Although a global effect of duration was found in both groups, the effect is likely attributable to different sources for each group. For example, the most poorly perceived category for CE listeners is the extreme (long) duration /r/ in initial position, receiving a rating approximately twice as poor as the original stimulus, while for the Japanese listeners, the same category was perceived approximately as accurately as its original counterpart. In general, the results suggest that the CE and Japanese listeners are sensitive to different aspects of the signal.

### 4. DISCUSSION

The results of the present experiment provide mixed support for the hypothesis that a sufficiently salient acoustic cue is capable of over-riding the attentional focus produced by a listener's native language phonological system. Some caution is necessary before completely accepting this conclusion, however. First, results from the CE listeners suggest that the duration manipulation may have produced slightly degraded stimuli. Despite the fact that the stimuli were identified perfectly by the CE listeners, their quality ratings indicate that the duration changes may have also changed the acoustic structure in other, more subtle ways. Thus, the methodology itself may require further investigation.

A second qualification of the present results is that only a small sample of Japanese listeners were used. The problem with small samples is that additional listeners can change the pattern of results. The two subjects who participated in the present experiment performed consistently across the different types of stimulus categories. Nevertheless, additional subjects would be useful to reduce the variability in the data even further.

Setting aside the above mentioned caveats, the present results do suggest that simple laboratory procedures may not provide a panacea for difficulties that listeners may have in perceiving certain phonemes. As Logan et al. (1991) noted, substantial amounts of training with naturally produced stimuli from multiple talkers may be the only way to reliably facilitate the perception of difficult-to-perceive nonnative phonemic categories.

### 5. REFERENCES

- Dissosway-Huff, P., Port, R.F., & Pisoni, D.B. (1982). *Research on speech perception: Progress Report No. 8*. Bloomington, IN: Indiana University.
- Goto, H. (1971). *Neuropsychologia*, **9**, 317-323
- Jamieson, D.G., Ramji, K.V., Neary, T.M. & Baxter, T.A. (1990). *Canadian Speech Research Environment (3.0)—User's manual*.
- Lively, S.E., Logan, J.S., & Pisoni, D.B. (1993). *JASA*, **94**, 1242-1255.
- Logan, J.S., Lively, S.E., & Pisoni, D.B. (1991). *JASA*, **89**, 874-886.
- MacKain, K., Best, C.T., & Strange, W. (1981). *Applied Psycholinguistics*, **2**, 369-390.
- Sheldon, A. & Strange, W. (1982). *Applied Psycholinguistics*, **3**, 243-261.

### 6. ACKNOWLEDGMENTS

Work supported by a grant from NSERC to JSL.

# ÉTUDE ACOUSTIQUE COMPARATIVE DES TONS "hoi" - "nga" EN VIETNAMIEN DU NORD ET DU SUD

Nguyen Ngoc-Quang

Université du Québec à Montréal, Département de linguistique  
C.P. 8888, succ. A, Montréal, Québec CANADA H3C 3P8

## 1. INTRODUCTION

Dans le système d'écriture en vietnamien moderne, il y a six signes diacritiques qui représentent six tons distincts de la langue. La principale distinction se manifeste aux modifications des hauteurs dans les courbes tonales. Selon les régions géographiques, certains tons dans ce système hexa-tonal diffèrent de l'un à l'autre [1, 2, 3].

Sur le plan phonologique, le vietnamien du nord possède six tons phonologiques tandis qu'en vietnamien du sud, il n'y en a que cinq [4, 5, 6]. Cette différence de nombre résulte du processus de fusionnement des deux tons "hoi" - "nga" que nous appelons ton 4 et ton 5 respectivement.

Notre étude vise à décrire et à caractériser, sur le plan de la réalisation phonétique, les propriétés acoustiques (les courbes tonales) des deux tons "hoi" - "nga" en vietnamien du nord et du sud. La normalisation de la fréquence fondamentale (F0) permet enfin de comparer ces deux tons en question.

## 2. MÉTHODOLOGIE

Le corpus d'analyse contient 54 mots monosyllabiques tirés de deux enregistrements réalisés au laboratoire de phonétique de l'Université du Québec à Montréal. Ces mots, enregistrés dans des phrases porteuses, contiennent chacune des neuf voyelles du vietnamien dans trois structures syllabiques (V, CV et CVC) ainsi que les deux tons "hoi" - "nga" (C1 = [s] et C2 = [m]). Deux locuteurs de sexe masculin, l'un du nord et l'autre du sud, appartenant à la classe moyenne, ont participé à l'étude.

Le corpus a été ensuite numérisé à 10 kHz à partir du logiciel d'analyse Speech Station. Un extracteur des valeurs de la fréquence fondamentale a été utilisé pour générer les valeurs de F0 de la rime, soit V ou VC [Figures 1, 2]. Les durées de ces segments ont été d'abord converties en pourcentage (100%) puis les valeurs de F0 ont été mesurées dans chaque intervalle de 5% de la durée convertie [7].

La compilation des données comporte deux étapes. La première étape consiste à trouver les courbes tonales caractéristiques de chaque type de syllabe en faisant la moyenne des valeurs de F0 des neuf voyelles [Tableaux 1 & 2]. La deuxième étape, quant à elle, consiste à trouver la fréquence fondamentale normalisée de chaque ton en faisant la moyenne des valeurs de F0 des trois types de syllabes [Tableau 3].

## 3. RESULTATS

La normalisation de la courbe tonale des tons 4 et 5 produit par les locuteurs de chaque région [Figure 3] nous permet de ressortir les observations suivantes:

@ Deux facteurs qui jouent un rôle décisif dans la caractérisation des tons sont la hauteur et le mouvement relatifs de F0.

@ Les tons 4 et 5 du vietnamien du sud ont une courbe tonale presque identique i.e. ils ont une même hauteur ainsi qu'un mouvement relatif analogue, descendant-montant.

@ les courbes tonales des tons 4 et 5 du vietnamien du nord ont cependant un mouvement quasi-ressemblant, descendant-montant-descendant, mais leurs hauteurs sont tout à fait distinctes entre elles; l'écart maximal de la hauteur du ton 5 est 49 Hz.

## 4. CONCLUSION

Nous avons fait ressortir l'importance de deux facteurs -la hauteur et le mouvement relatifs- que sont à la base de la différence entre les tons "hoi" - "nga" de deux régions dialectales du Vietnam.

Nous retenons également que, sur le plan phonologique, les tons "hoi" - "nga" du vietnamien du sud ont subi un processus de fusionnement de la hauteur relative et ils ont une même réalisation phonétique, un mouvement descendant-montant. Comparativement au vietnamien du nord, il apparaît que la neutralisation consiste à une perte de la hauteur relative du ton "nga".

Cependant, les tons 4 et 5 du vietnamien du nord gardent précieusement, chacun de leur part, leurs propriétés acoustiques particulières, descendant-montant-descendant et une distinction de hauteur relative.

Il serait sans doute intéressant de chercher les indices psycho-acoustiques, par le biais des tests perceptifs, qui permettent de différencier ces deux tons "hoi" - "nga" du vietnamien du nord.

## 5. BIBLIOGRAPHIES

[1] LE, V.L. (1948). *Le parler vietnamien*. Paris.

[2] BULTEAU, R. (1953). *Cours d'annamite*. Paris.

[3] DONALDSON, J. (1963). A study of the "nang" tone in the northern vietnamese. Van Hoa Nguyet San 12, pp. 1151-4.

[4] NGUYEN, D.H. (1987). Vietnamese. The World's Major Languages, pp. 777-795, Oxford University Press, New York.

[5] Thompson, L. (1959). Saigon Phonemics. Journal of the Linguistic Society of America, vol. 35, pp. 454-476, Waverly Press, Baltimore.

[6] HAUDRICOURT, A.G. (1954). De l'origine des tons en vietnamien. Journal Asiatique, vol.242, pp.69-82, Paris.

[7] HOWIE, J.M. (1970). The Vowels and Tones of Mandarin Chinese: Acoustical Measurements and Experiments. Ph.D. dissertation, Indiana University.

F0	Ton 4			Ton 5		
	V	CV	CVC	V	CV	CVC
Durée en %						
0	116	112	119	121	114	110
5	114	108	114	120	111	106
10	111	106	112	119	108	104
15	108	104	110	116	107	103
20	105	102	108	115	106	103
25	103	100	106	115	106	104
30	102	99	105	117	108	106
35	101	100	104	119	111	110
40	102	101	102	125	119	115
45	105	104	104	133	130	122
50	111	112	112	140	142	129
55	120	128	125	163	154	148
60	134	149	139	188	166	166
65	150	166	154	214	197	196
70	169	184	178	244	217	216
75	188	203	202	266	237	233
80	217	229	221	282	254	245
85	230	241	235	294	272	249
90	236	247	239	296	277	250
95	233	247	237	288	276	249
100	224	245	230	277	274	246

Tableau 1: Les fréquences fondamentales des tons 4 et 5 du vietnamien du nord associées à trois types de syllabes

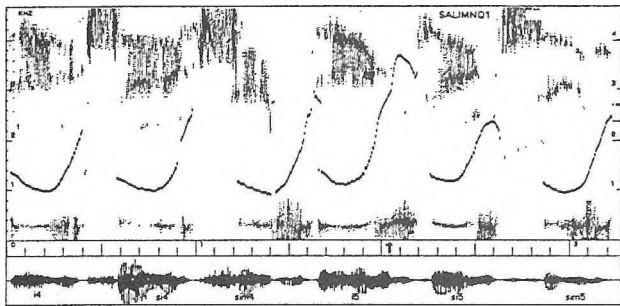


Figure 1: Les courbes tonales des tons 4 et 5 dans les mots [i, si, sim] produits par le locuteur du vietnamien du nord

F0	Ton 4			Ton 5		
	V	CV	CVC	V	CV	CVC
Durée en %						
0	111	118	121	111	123	124
5	110	111	114	109	116	116
10	107	107	110	107	110	111
15	105	103	106	105	104	107
20	104	100	103	104	101	104
25	103	100	102	103	101	103
30	104	102	103	105	102	104
35	105	104	104	106	104	106
40	108	107	106	109	106	107
45	112	112	109	113	110	109
50	117	118	115	119	116	114
55	125	126	123	126	123	128
60	136	138	136	137	135	140
65	149	151	148	150	145	154
70	160	165	163	162	159	170
75	173	176	180	175	172	181
80	183	186	190	187	182	187
85	191	189	196	195	186	192
90	193	187	199	196	188	192
95	192	185	200	193	187	191
100	188	184	202	189	186	187

Tableau 2: Les fréquences fondamentales des tons 4 et 5 du vietnamien du sud associées à trois types de syllabes

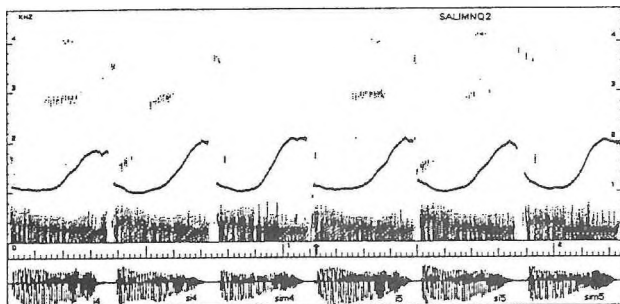


Figure 2: Les courbes tonales des tons 4 et 5 dans les mots [i, si, sim] produits par le locuteur du vietnamien du sud

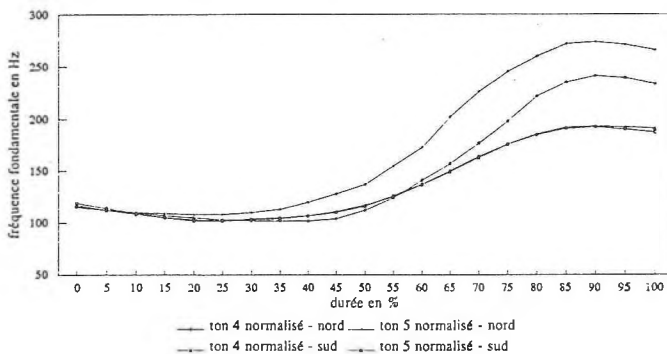


Figure 3: Les tons "hoi" & "nga" du vietnamien du nord et du sud

F0	nord		sud	
	Ton 4	Ton 5	Ton 4	Ton 5
Durée en %				
0	116	115	117	119
5	112	112	112	114
10	110	110	108	109
15	107	109	105	105
20	105	108	102	103
25	103	108	102	102
30	102	110	103	104
35	102	113	104	105
40	102	120	107	107
45	104	128	111	111
50	112	137	117	116
55	124	155	125	126
60	141	173	127	137
65	157	202	149	150
70	177	226	163	164
75	198	245	176	176
80	222	260	186	185
85	235	272	192	191
90	241	274	193	192
95	239	271	192	190
100	233	266	191	187

Tableau 3: Les fréquences fondamentales normalisées des tons "hoi" & "nga" du vietnamien du nord et du sud

# Relâchement et abrègement des voyelles: phonétique ou phonologie?

François Poiré

Université du Québec à Montréal, Département de linguistique  
C.P. 8888, succ. A, Montréal, Québec (Canada) H3C 3P8

## 0. Introduction

Nous voulons dans ce travail jeter un nouveau regard sur la relation qui existe entre le relâchement des voyelles hautes et leur abrègement en français du Québec. Nous allons arguer que le relâchement est un phénomène phonétique qui semble toucher toutes les voyelles, au même titre que l'abrègement et l'allongement de ces segments.

## 1. Problématique

Dans la section sur les lois de durée de Marchal (1980), on peut lire:

«(...) les voyelles (...) dans une syllabe accentuée fermée par /p/, /t/ ou /k/ (...) deviennent très brèves (on pourra souvent observer une tendance à l'ouverture des voyelles hautes.)» p.84

D'autre part, il ajoute à la page 163, lorsqu'il aborde plus spécifiquement le relâchement des voyelles hautes:

«Ce phénomène n'est pas gênant pour la communication tant qu'il n'y a pas réduction d'une opposition de timbre.»

Il semble donc que d'un côté, le phénomène du relâchement soit un phénomène régulier en français du Québec, et d'un autre côté, on ne l'associe qu'aux voyelles hautes en raison d'une interférence communicationnelle causée par une opposition vocalique menacée /i, e/.

De plus, s'il y a relâchement, c'est en raison de la présence en position post-vocalique de certaines consonnes dites abrègées. S'il s'agit de consonnes dites allongées, on ne parle plus de relâchement. Il y aurait alors distribution complémentaire parfaite. Le tableau qui suit synthétise ce que Marchal (1980) dit explicitement au sujet de cette distribution concernant l'effet des consonnes sur les voyelles hautes.

	relâchement	durée
occlusives sourdes	oui	abrè.
occlusives sonores	non	all.
fricatives sourdes	oui	abrè.
fricatives sonores	non	all.
nasales	?	?
l	?	?
r	non	all.

Tableau 1. Effet des consonnes entravantes sur les voyelles hautes selon Marchal 1980 (abrè. = abrègement, all. = allongement).

La distribution complémentaire semble parfaite. Elle le demeurera jusqu'à ce que l'effet des nasales et de /l/ soit élucidé.

La question qui nous intéresse est donc la suivante: le relâchement des voyelles est-il une conséquence normale de l'abrègement (ou bien: l'allongement préserve-t-il du relâchement)? Si oui, peut-on alors parler d'un phénomène phonologique?

## 2. Méthodologie

Pour tenter de répondre à ces questions, nous avons construit un corpus dans lequel les voyelles du français /i, E, a/ se retrouvent en syllabe ouverte ou fermée par les consonnes /p, b, f, v, l, r, m/. Chaque mot a été enregistré trois fois à tempo normal dans la phrase porteuse «répète x deux fois» par un seul locuteur. Chaque phrase a été numérisée à 10k par échantillon pour qu'ensuite soient extraites les durées des voyelles ainsi que les valeurs de F1 et F2 en leur centre, à l'aide du logiciel Speech Station. Nous ne parlerons ici que de durée et de F1.

Afin de comparer les valeurs de durée et de formants, les voyelles en syllabe ouverte servent de point de référence. Ainsi, la durée de /i/ en syllabe non entravée vaut 100%. Si une consonne a un effet allongeant de 5% sur cette voyelle, nous dirons que cet effet vaut +5. Tous les tableaux de la section suivante suivent ce principe.

## 3. Résultats

### 3.1 La durée

Le tableau 2 montre l'effet des sept consonnes du corpus sur la durée des trois voyelles. Les consonnes sont ordonnées non pas selon les classes naturelles mais selon l'importance de l'effet, de la plus grande valeur négative à la plus grande valeur positive.

	i	E	a
p	-35.01	-31.33	-33.05
f	-22.04	-07.45	-16.97
m	-13.18	+09.00	+17.50
b	-03.06	+19.02	+19.95
l	+13.42	+40.84	+47.45
v	+86.52	+55.09	+58.34
r	+95.43	+88.38	+72.98

Tableau 2. Effet des consonnes sur la durée des voyelles exprimé en %.

Les valeurs soulignées montrent le passage de l'abrègement à l'allongement. Ce tableau montre que les trois voyelles subissent les mêmes effets quant à leur durée. La seule différence notable concerne les effets de /m, b/ sur le /i/.

Nous sommes donc en présence d'une hiérarchie où les valeurs changent de manière graduelle et non catégorielle

### 3.2 Les valeurs de F1

Le relâchement est habituellement associé à l'ouverture des voyelles même si certains auteurs essaient encore de lier le phénomène à une baisse de la force articulatoire (Ostiguy et Tousignant, 1993). En suivant Delattre (1951), l'augmentation des valeurs de F1 serait la conséquence acoustique de l'ouverture de la cavité buccale. F2, toujours selon le même auteur, diminuerait à mesure que la cavité frontale s'allonge (à mesure que la langue recule).

Si le relâchement est réellement associé à l'abrègement, les valeurs de F1 des voyelles abrègées devraient augmenter.

Le tableau 3 montre les valeurs de F1 de nos trois voyelles sous l'influence des consonnes abrégées du tableau 2.

	i	E	a
p	+08.19	+01.10	+09.62
f	+04.92	+03.29	+08.87
m	+09.83	+06.68	0.00
b	+08.18	-01.10	+10.48

Tableau 3. Effet des consonnes sur le F1 des voyelles exprimé en %.

On peut voir que les valeurs négatives du tableau 2 sont corrélées à des valeurs positives du tableau 3. Mais pour que cette corrélation soit concluante, les valeurs positives de durée du tableau 2 devraient être associées à des valeurs négatives de F1. Le tableau 4 donne ces valeurs d'effet sur F1 des consonnes allongées du tableau 2.

	i	E	a
l	+08.19	-01.10	+05.65
v	0.00	0.00	0.00
r	-06.56	+24.17	+09.68

Tableau 4. Effet des consonnes sur le F1 des voyelles exprimé en %.

On retrouve dans le tableau 4 et des valeurs positives et des valeurs négatives. C'est donc dire que la corrélation n'est pas présente.

#### 4. Discussion

Nos résultats montrent qu'il est très difficile d'établir une corrélation directe entre l'abrègement des voyelles et le relâchement. De plus, l'allongement ne semble pas préserver ces mêmes voyelles des changements de valeur de F1. Ces deux points

vont à l'encontre de ce qui est dit habituellement sur ces phénomènes. Autrement dit, les effets des consonnes sur le F1 des voyelles (tableaux 2 et 3) ne permettent pas de dégager deux ensembles de données distincts justifiant la présence d'une règle de relâchement au niveau de la phonologie. Seul le comportement de /i/ en ce qui concerne sa durée (tableau 1) va dans le sens d'une distinction phonologique au niveau des classes de voyelles: si ce segment est plus facilement affecté dans sa durée par les consonnes entravantes, c'est qu'il appartient à une classe de voyelles dites ultra légères (Cedergren et Simoneau, 1985).

#### 5. Conclusion

Que F1 soit ou non le corrélant acoustique du relâchement, nos données montrent que les voyelles subissent presque toujours les mêmes modifications formantiques selon la consonnes entravantes. Le relâchement serait alors un simple phénomène phonétique général du français, du moins dans sa version parlée au Québec. Il aurait obtenu son statut de vedette de la prononciation de cette variété en ce qui concerne les voyelles hautes en raison de la distribution de l'espace phonologique de notre système vocalique en interaction avec les particularités de la réalisation phonétique propres à cette variété.

#### 6. Bibliographie

- Carton, Fernand (1974) Introduction à la phonétique de français, Paris, Bordas.
- Cedergren, H. et L. Simoneau (1985) "La chute des voyelles hautes en français de Montréal 'As-tu entendu la belle syncope?'" , Les tendances dynamiques du français parlé à Montréal, tome 1, ss dir. M. Lemieux et H. Cedergren, Québec, Office de la langue française.
- Delattre, Pierre (1951) The Physiological Interpretation of Sound Spectograms, PMLA LXVI, repris in Delattre (1966), Studies in French and Comparative Phonetics, Mouton and CO, The Hague.
- Marchai, Alain (1980) Les sons et la parole, Montréal, Guérin.
- Ostiguy, I. et C. Tousignant (1993) Le français québécois, Normes et usages, Montréal, Guérin.
- Walter, Henriette (1977) La phonologie du français, Paris, Presses universitaires de France, coll. le linguiste.

# Le focus et l'intonation en français parlé à Montréal

Louise Levac  
Université du Québec à Montréal, Département de linguistique  
C.P. 8888, succ. A, Montréal, Québec (Canada) H3C 3P8

## INTRODUCTION

Notre recherche comporte deux volets. D'une part, nous voulons connaître quelles sont les caractéristiques linguistiques du focus et ce qui le distingue comme catégorie prosodique des autres catégories comme, par exemple, l'intonation terminale, l'intonation continuative, etc. D'autre part, nous voulons savoir si les variations de la fréquence fondamentale expliquent le fait que les auditeurs québécois ont l'habileté de discriminer un focus des autres types catégoriels.

## 1. METHODOLOGIE

Notre corpus est constitué de quatre extraits d'environ cinq minutes puisés dans l'enregistrement de quatre locuteurs (deux femmes et deux hommes) du *Corpus Sankoff-Cedergren* (Le français de Montréal, 1971). Ces extraits sont filtrés et numérisés à 8000 sur un Texas Instrument. Ils sont analysés de deux façons: par une segmentation perceptuelle et par une segmentation acoustique.

### 1.1 Segmentation perceptuelle

Deux locutrices québécoises ont segmenté à l'écoute les enregistrements en quatre catégories prosodiques: soit, 1) l'intonation complète, 2) l'intonation partielle, 3) l'accentuation secondaire et 4) le focus. Cette catégorisation prosodique s'est faite en deux séances; la première pour identifier les trois premières catégories et la seconde pour identifier le focus. Avant la séance d'écoute, des critères impressionnistes généraux ont été remis et la consigne a été de marquer distinctivement sur un texte orthographié de l'enregistrement les endroits où elles percevaient une catégorie. Le tableau 1 ci-haut présente ces critères que nous avons élaborés et le tableau 2 fournit des exemples de segmentation perceptuelle.

### 1.2 Segmentation acoustique

La durée en ms de la voyelle marquée d'une marque prosodique et la valeur en Hz de son fondamental sont calculées et comparées entre elles dans le but de faire ressortir les différences acoustiques générales qui opposeraient le focus aux autres catégories prosodiques. La segmentation de la voyelle se fait prioritairement sur le signal. Le changement dans l'onde oscillographique sert à fixer les bornes initiales et finale du segment. A l'écoute, nous décidons de l'étiquette phonétique du segment. A l'aide de l'extracteur du programme *Speech Station*, les valeurs du fondamental pour un énoncé sont calculées automatiquement et affichées sous la forme d'un contour intonatif. La valeur en Hz au 23 de la voyelle est retenue pour l'analyse

---

## TABLEAU 1: Méthode d'écoute et critères impressionnistes

---

Première séance: Ecouter tout l'extrait d'un locuteur, puis à la réécoute et en vous servant du texte écrit

1) Marquer d'un ] une intonation complète. (Mettre dans cette catégorie les intonations suivies d'une pause.)

2) Marquer d'une } une intonation partielle, interne à l'intonation complète. (Mettre dans cette catégorie les hésitations qui ne sont pas suivies d'une pause.)

3) Marquer d'une ) une prééminence accentuelle secondaire, qui est interne à l'intonation complète ou à l'intonation partielle.

---

Deuxième séance: Ecouter tout l'extrait d'un locuteur, puis à la réécoute et en vous servant du texte écrit

1) Marquer d'un + une prééminence accentuelle (très) forte et (très) haute.

---

---

## TABLEAU 2: Exemples d'énoncés segmentés à l'écoute (Les marques entre parenthèses indiquent que les juges ont des interprétations variées.)

2.23 C'est mon frère } qui l'a acheté } quand elle est morte.] Elle est morte}]+ au mois ( ) de février.]

7.2 Puis ensuite ] l'école a été changée } pour Basile-) Routhier.]+

70.8 Il faut pas } que je m'étende } (+) là-dessus]+

---

comparative des différentes marques prosodiques. Le mode de comparaison des valeurs en Hz est le demi-ton, il est calculé en mettant en rapport la valeur de la voyelle marquée prosodiquement à celle de la voyelle adjacente de gauche. Le tableau 3, à la page 2, donne un exemple des valeurs obtenues par la segmentation acoustique.

## 2.RESULTATS

### 2.1 Analyse linguistique

Les résultats de la segmentation perceptuelle montrent que le focus peut être entendu indépendamment des autres catégories, mais qu'il peut aussi être cooccurrent aux catégories ] et ]. Il est optionnel: de grandes parties des extraits ne contiennent aucune marque de focus alors que toutes les phrases des extraits comportent au moins la marque ].

Une analyse discursive des lieux d'ancrage des marques du focus montre qu'elles peuvent être interprétées comme des marqueurs prosodiques qui ont une fonction déictique: le focus fait ressortir un item énoncé de façon à y attirer l'attention de l'interlocuteur; il s'agit d'un marqueur extralinguistique qui oriente l'attention sur des sujets d'intérêt pour l'un ou l'autre des interlocuteurs. On peut paraphraser cette fonction par les phrases suivantes: "Ceci, ça m'intéresse d'en parler" et "Cela, ça t'intéresse d'en parler".

### 2.2 Analyse perceptuelle

La mise en commun du travail d'écoute de l'une et l'autre des auditrices indique qu'il existe un taux d'accord moyen de 55.77 % entre les catégories prosodiques distinguées et les syllabes marquées. Il existe un taux d'accord partiel moyen de 11.81 %, c'est-à-dire que les locutrices ont toutes les deux perçu une marque prosodique mais qu'elles ne l'ont pas identifiée par la même étiquette. Le taux de désaccord moyen est de 32.42 %; dans ces cas, une seule des locutrices a perçu une marque prosodique et l'a catégorisée.

### 2.3 Analyse acoustique

Nous présentons les résultats préliminaires de l'analyse acoustique pour lesquelles les juges ont un taux d'accord. Le tableau 3 présente les deux paramètres acoustiques (durée et pitch) et le rapport en demi-tons. Dans l'énoncé 2.6 du locuteur 2, *Il n'y avait pas de transports*, la dernière syllabe du mot *transports* a été perçue comme portant une marque d'intonation complète ] et une marque de focus +. La voyelle [ɔ] de cette syllabe a une durée de 48.9ms et un pitch de 135.1 Hz. Le rapport des valeurs du fondamental entre la pénultième [ʒ] et la dernière syllabe est de 3.73 demi-tons.

## CONCLUSION

Dans la parole spontanée, les marques prosodiques sont entendues et catégorisées par les locuteurs natifs. Avec des critères impressionnistes et des consignes d'écoute, nous avons obtenu des données qui montrent l'existence de notions linguistiques comme l'intonation complète, l'intonation partielle, la proéminence

secondaire, le focus. Dans notre recherche, le focus a été critérié comme une proéminence forte et haute. Les résultats des données d'écoute montrent que le focus est une marque distincte des autres marques prosodiques, qu'il est optionnel et qu'il est cooccurrent ou indépendant. Les résultats de l'analyse acoustique mettront en lumière les corrélats physiques propres à chacune des catégories prosodiques.

---

TABLEAU 3: Exemple d'analyse acoustique

2.6 Il n'y avait pas ( ) de transports.]+

étiquette i a v a p a t r ǎ s p ɔ w

durée 48.9

Hz 109 135.1

---

## BIBLIOGRAPHIE

- Corpus Sankoff-Cedergren (1971) H.J. Cedergren. Département de linguistique, UQAM.
- Cedergren *et al.* (1990) *L'accentuation québécoise : une approche tonale*. Revue québécoise de linguistique, vol. 19, no. 2, p. 25-38.
- Levac, L. (1993) *Deictic prosodic Focus*. Poster présenté à NWAV 22, Université d'Ottawa, Ottawa, Ontario, (Octobre 1993).