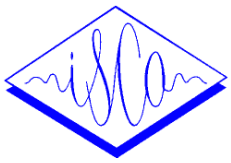
ISCA Archive
http://www.isca-speech.org/archive



ETRW on Sound Patterns of
Spontaneous Speech (SPoSS)
La Baume-les-Aix, France
September 24-26, 1998

LANGUAGE DEPENDENT AND INDEPENDENT SPONTANEOUS SPEECH PHENOMENA

Tzu-Ting SU and Patricia BASSET
Institut de Phonétique, I.L.P.G.A. 19, rue des Bernardins, 75005 Paris
e-mail : patriciapier@magic. fr

RESUME

Cet article présente des résultats expérimentaux sur la disparition de 372 phonèmes du français spontané et de 175 phonèmes du mandarin spontané de Taiwan. Quatre points ont été examinés. Premièrement, le pourcentage de la disparition des phonèmes est le même dans les deux langues: 3.5%. Deuxièmement, le phénomène de disparition concerne essentiellement les voyelles en français et les consonnes en mandarin taïwanais. Par conséquent, les groupes de consonnes sont fréquents et les hiatus rares en français, alors qu'en mandarin taïwanais, les tendances sont inverses. Troisièmement, la disparition se produit principalement sur les mots grammaticaux dans les deux langues: 75% de phonèmes disparus appartiennent à des mot grammaticaux. Enfin, la disparition des consonnes a surtout lieu en position coda en français mais seulement en position initiale en mandarin taïwanais (toutes les consonnes, sauf deux nasales, ne peuvent se trouver qu'en position initiale).

1. INTRODUCTION

There is growing interest for spontaneous versus read speech. Most studies concern Indo-European languages and one language at a time, like English [1], German [2] and French [3]; no unified method is used, so it is difficult to compare the results. Consequently, in this study, the same methodology is applied to investigate the disappearance of phonemes in spontaneous speech for two languages genetically unrelated: French (Romance language) and Mandarin Chinese of Taiwan (Sino-Tibetan language).
The investigation is based on spectrographic and perceptive analysis.
The same points were investigated for both languages in this preliminary study. First, the percentage of deletion of phonemes. Second, the percentage of deletion of vowels versus deletion of consonants. Third, the relationship between the deletion of phonemes and the class of the words (lexical or grammatical). Finally the relationship between the deleted phonemes and their position in a syllable and in an utterance.

2. METHOD

The corpus consists of ten minutes of spontaneous mandarin uttered by two male speakers native of Taiwanese Mandarin, and 16 minutes of spontaneous French uttered by two French male speakers, in an anechoic chamber.
The 26 minutes of speech were orthographically transcribed. Signals (with visual help of spectrograms) were segmented, phonetically transcribed (whenever possible) and prosodically annotated on a computer. The criterion of disappeance is based on spectrographic analysis. Some problems of transcription occured when phonemes were invisible on the spectrogram but were audible when presented to listeners with their surrounding phonemes. This problem reflects the fact that a phoneme not visible on the spectrogram may leave somewhere a trace, as noted by Manuel [4]: "lenitions are less complete than they first appear". Table 1 represents the four categories with which a phoneme can be characterised according to whether it is visible on spectrogram and it is perceived. Therefore we prepared a perception test for "invisible" phonemes. Two listeners (two for Mandarin and two for French) had to transcribe what they heard when excerpts of speech corresponding to the phonemes surrounding the supposedly deleted one were presented. Following these perception tests we consider a phoneme as deleted if it is invisible on the spectrogram and not perceived by listeners (category 4 on the Table below).

	1	2	3	4
visible on spectrogram	+	-	+	-
perceived	+	+	-	-

Table 1. The four categories adopted to classify phonemes

3. RESULTS AND DISCUSSIONS

3-1. Percentage of deleted phonemes

Table 2 indicates that the percentage is surprisingly the same for both languages: 3,5%. This percentage is low. Further study with more languages and speakers are needed to verify such a low percentage of deletion.

French			Taiwanese Mandarin		
total	deleted	%	total	deleted	%
10558	372	3.5	4780	175	3.5

Table 2. Total versus deleted phonemes

3-2. Vowels versus consonants

Table 3 shows the number of vowels and consonants which were deleted.
Vowels are more likely to be deleted than consonants in French (5.3% of vowels versus 1.9% of consonants deleted).
Consonants are more likely to be deleted than vowels in Taiwanese Mandarin (4.4% of consonants versus 2.9% of vowels deleted).

French				Taiwanese Mandarin		
	total	deleted	%	total	deleted	%
V	4933	263	5.3	2307	67	2.9
C	5288	103	1.9	1766	77	4.4
G	337	6	1.7	707	22	3.1

Table 3. Total versus deleted phonemes for V (vowel), C (consonant), and G (glide)

Fig. 1 illustrates a typical example of vowel deletion in French and of consonant deletion in Taiwanese Mandarin. In the word "de" [də] of [**tudəsɥit**] "tout de suite" (right now) in French, [ə] disappeared, while for the word [tə] of [**cçɥétə**] (think) in Taiwanese Mandarin, [t] disappeared.

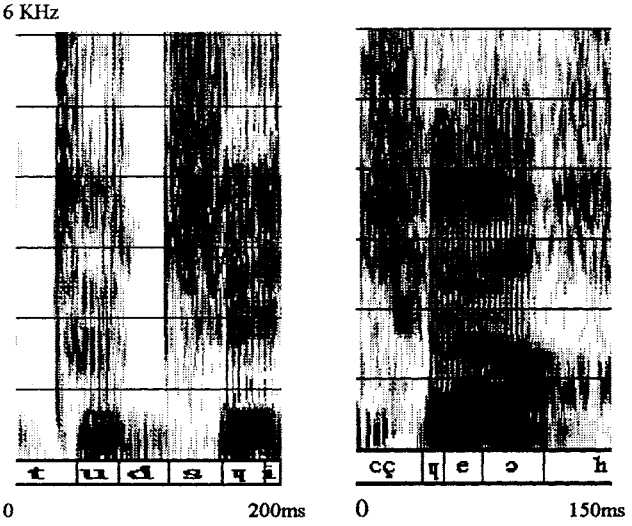


*Figure 1: [**tud(ə)sɥit**] "tout de suite" (right now) in French and [**cçɥe(t)ə**] (think) in Taiwanese Mandarin*

As far as our results are concerned, the deletion of vowels in French leads to a great number of clusters and hiatus are not frequent, while in Taiwanese Mandarin, the formation of clusters is rare but succession of vowels is more frequent.

3-2-1. Vowels

Table 4 illustrates which vowels are deleted in both languages. Some points can be observed:
1) Deletion of the central vowel schwa is very frequent in French. Such a phenomenon is well documented [5]. In our data, 29.5% of schwa disappear. Similarly, in Taiwanese Mandarin, the schwa, nasalised or not, disappears (16%).
2) In Taiwanese Mandarin, oral close vowels tend to disappear. Thus, the tendency that close vowels are less resistant than open vowels [6] seems quite systematic in Taiwanese Mandarin. In French, all vowels may disappear, even the open vowel [a] (1.7%). 61.4% of deleted [a] belong to the word [**paʁsk**] "parce que" (because), a very frequent word that Bally [7] would call "worn out". In our data, two productions were found : [**pask**] and [**psk**]. We follow the suggestion made by Ohala [8]: more than one entry may exist for very frequent words.
3) Note that [**i**] in Taiwanese Mandarin is at the same time central and close. This vowel is the most deleted one (12.1%).
4) The nasalised vowels [**ən**] (nasalised schwa) and [en] in Taiwanese Mandarin disappear only when they are preceded or followed by a nasal consonant, which has a strong vocalic character (noting that the only disappeared [**en**] is preceded by [**m**] and [**j**]) For example: [**wǒmən**] (we) becomes [**wǒm**]. In French, 62% of deleted nasal vowel [**ɑ̃**] occurs in the frequent word [**ɑ̃fɛ̃**] "enfin" (finally). Two productions were found : [**ɑ̃fɛ̃**] and [**fɛ̃**].

It would be interesting to examine more tone languages to test whether the fact that vowels are carriers of tones may lead to less frequent deletion of vowels.

French			Deleted / Total (%)
i 29/666(4.5)	**y** 7/226(3.1)		**u** 4/153(2.6)
e 2/424(0.5)	**ø** 3/45(6.6)		**o** 3/158(1.9)
ɛ 22/758(2.9)	**œ** 0/123	**ə** 158/536(29.5)	**ɔ** 0/306
		a 13/811(1.7)	
ɛ̃ 0/184		**ɑ̃** 10/358(2.7)	**ɔ̃** 0/54

Taiwanese Mandarin			Deleted / Total (%)
i 5/190(2.6)	**y** 1/51(2)	**i** 26/214(12.1)	**u** 2/97(2.6)
e 0/95		**ə** 13/291(4.5)	**o** 0/231
		ɚ 0/6	
		a 0/104	
in/ŋ 0/87		**ən** 18/156(11.5)	**oŋ** 0/63
en 1/119(0.8)		**an/ŋ** 0/183	
ei 0/96	**ai** 0/104	**au** 0/79	**ou** 0/141

Table 4. Deleted vowels versus total number of vowels (%) in French and Taiwanese Mandarin

3-2-2. Glides

Table 5 shows the deleted versus total number of glides. The [j] (2.3%) and [w] (4.3%) disappeared in Taiwanese Mandarin but not in French, where only the [ɥ] disappeared (9.8%). For example: [pɥi] "puis" (then) becomes [pi].

	French	Taiwanese Mandarin
j	0/162	8/346 (2.3%)
w	0/55	13/304 (4.3%)
ɥ	6/61 (9.8%)	1/57 (1.8%)

Table 5. Deleted glides versus total number of glides in French and Taiwanese Mandarin

3-2-3. Consonants

The results of deleted versus total number of consonants are shown in Table 6. Some points can be noticed:
1) In French, only voiced fricatives ([v]: 4.6%, [ʁ]: 4.3%) and lateral approximant phoneme [l] (10.8%) disappeared, while in Taiwanese Mandarin the deletion phenomenon is quite scattered.
2) Voiced fricatives disappeared in both languages (ʐ: 15.3% in Taiwanese Mandarin). It is known that voiced fricatives have weaker frication than unvoiced fricatives because of aerodynamic constraints [9] (conflict between frication and voicing). As a result, they are more likely to be deleted.
3) Regardless of the scattered tendency of consonants in Taiwanese Mandarin, surprisingly none of labial consonants disappeared. Straka [10] concluded that the labial [p] is less likely to disappear than the dental [t] and velar [k].
4) In Taiwanese Mandarin the most deleted consonant is [x] (20.3% disappeared).

The deletion phenomenon of consonant was also mentioned by Lien [11] In Taiwan Southern Min, a tone language as well, where consonants such as voiceless stop aspirated [k'] can disappear. For example, [lipk'i] (enter in) becomes [lipi].

French

		labial	dental	palatal	velar	uvular	Deleted / Total (%)
plosive	voiceless	p 0/370	t 0/485		k 0/526		
	voiced	b 0/147	d 0/444		g 0/82		
fricative	voiceless	f 0/139	s 0/592	ʃ 0/71			
	voiced	v 8/174 (4.6)	z 0/115	ʒ 0/177		ʁ 28/647(4.3)	
nasal		m 0/365	n 0/227				
lateral			l 66/609 (10.8)				

Taiwanese Mandarin

		labial	dental	retroflex	palatal	velar	Deleted / Total (%)
plosive	not aspirated	p 0/91	t 13/277 (4.7)			k 7/114 (6.1)	
	aspirated	ph 0/9	th 1/61 (1.6)			kh 0/43	
affricate	not aspirated		ts 0/61	tʂ 3/118 (2.5)	cç 8/103 (7.8)		
	aspirated		tsh 0/7	tʂh 1/26 (3.8)	cçh 1/57 (1.8)		
fricative	voiceless	f 0/38	s 1/46 (2.2)	ʂ 10/232 (4.2)	ç 0/120	x 21/103 (20.3)	
	voiced			ʐ 6/39 (15.3)			
nasal		m 0/63	n 1/95 (1.1)				
lateral			l 4/63 (6.3)				

Table 6. Deleted consonants versus total number of consonants (%) in French and Taiwanese Mandarin

3-3. Lexical versus grammatical word

The relationship between the class of a word and the deletion of phonemes was investigated. It is well known that phonemes are more likely to be deleted in grammatical words than in lexical words [12].
We divided all the phonemes into two classes, depending on whether they belong to lexical or grammatical words.

For French, grammatical words are pronouns, articles, auxiliaries, conjunctions, and prepositions. For Taiwanese Mandarin, we add postpositions and particles. Nouns, adjectives, verbs and adverbs are lexical words.
Table 7 shows that most of the deleted phonemes belong to grammatical words in the two languages (75% in grammatical words versus 25% in lexical words). In addition, Table 8 shows that phonemes are more likely to disappear in grammatical words than in lexical words

in both languages (8% versus 1.3% in French and 5.5% versus 1.7% in Taiwanese Mandarin). Our results confirm therefore the previous studies [13].

	French		Taiwanese Mandarin	
Deleted	372	%	175	%
Grammatical	278	75	132	75
Lexical	94	25	43	25

Table 7. Deleted phonemes in grammatical words and in lexical words

	French			Taiwanese Mandarin		
	total	deleted	%	total	deleted	%
G	3445	278	8	2382	132	5.5
L	7113	94	1.3	2498	43	1.7

Table 8. Deleted phonemes in grammatical words (G) versus in lexical words (L)

3-4. Position in the syllable and in the utterance

Syllables onset are well known to be more resistant than coda [14]. Indeed, according to our data, in French, consonants in position coda are much more likely to be deleted (89% deleted in coda).
In Taiwanese Mandarin, the deleted consonants are in initial position, since the consonants (except nasal consonants [n] and [ŋ]) exist only at the onset of the syllable. As far as deleted [ən] is concerned in our data, the entire rhyme is deleted, which confirms that final nasal belongs to the vowel for Mandarin as Hagège and Haudricourt [15] suggested. However, in considering the position in an utterance (between two pauses) we found that in Taiwanese Mandarin no phoneme among 236 disappeared in the beginning of an utterance.

4. CONCLUSION

We have tried to examine in this paper how far deletion of phonemes in spontaneous speech depends on the particular language. The preliminary analysis shows that the percentage of deletion is about the same in French and Taiwanese Mandarin. The deletion of phonemes concerns essentially vowels in French, but consonants in Taiwanese Mandarin. The phonemes are more likely to be deleted in grammatical words than in lexical words for both languages.
Most of the deleted consonants in French are coda consonants, while in Taiwanese Mandarin all the deleted consonants are syllable onset consonants, since consonants can only be syllable initial (if we consider final nasal as part of the vowel).

5. ACKNOWLEDGEMENT

We would like to thank Professor Vaissière for her help without which this work would not have been achieved.

6. REFERENCES

[1] F. Nolan, "*Overview of English Connected Speech Processes*", AIPUK, Vol. 31, pp: 15-26, 1996.
[2] K Kohler. "*The phonetic realization of schwa syllables in German*", AIPUK; Vol 31., pp. 11-14, 1996.
[3] D. Duez, "*On spontaneous French speech: aspects of the reduction and contextuel assimilation of voiced stops*", Journal of Phonetics, Vol. 23, pp. 407-427, 1995.
[4] S.Y. Manuel, "*Speakers nasalise [ð] after [n] but listeners still hear [ð]*, Journal of Phonetics, Vol. 23, pp. 453-476, 1995.
[5] H. Walter; "*La dynamique des phonèmes dans le lexique français contemporain*", France-Expansion, pp. 284-285, 1976.
[6] [10] [14] G. Straka, "*Les sons et les mots*", Klincksieck, pp. 213-228, 1979.
[7] Ch. Bally, "*Le langage et la vie*", Payot, Paris, pp. 148, 1926, cité par Hagège C & Haudricourt. A; La phonologie panchronique, PUF, Paris, pp. 115, 1978.
[8] J. Ohala, Conference at ILPGA, Paris, 1994.
[9] J. Ohala, "*The relation between phonetics and phonology*", The Handbook of Phonetic Sciences, Blackwell, PP. 687-689, 1997.
[11] C. Lien, "*Studies on directional complement in Taiwan Southern Min - dialect typologie and historical studies*", Chinese language and linguistics IV: typological studies of languages in China, Academia Sinica, Vol. 2, Taipei, pp. 379-403, 1997.
[12] B. Malmberg, "*La phonétique*", PUF, Paris, pp. 113-114, 1954.
[13] C. Rehor and M. Pätzold, "*The phonetic realization of function words in German spontaneous speech*", AIPUK, Vol. 31, Kiel, pp. 5-10, 1996.
[15] C. Hagège, A. Haudricourt, "*La phonologie panchronique*", PUF, Paris, pp. 90, 1978.