Negative Transfers in Intonation between English and Japanese

June-ko Matsui
The University of Tsukuba

INTRODUCTION

Numerous factors account for the unnaturalness a speaker exhibits when talking in a second language: differences in vowel types, timing, pause, amount of voicing, pitch, and intensity contribute toward the sense of "foreignness" in the second-language speaker's utterances. Although the native speaker is able to discern that differences exist between the foreign and native speech, it is often unclear exactly which factors are different, and to what degree. This paper attempts to clarify which phonetic negative transfers contribute largely to the sense of foreignness in second language speakers, and how prosodic features such as intensity, pitch, length, amount of voiced sounds, and pause, are related to each other.

METHODOLOGY

Nine native American speakers and nine native Japanese speakers were asked to read a list of exclamatory sentences in both English and Japanese. The utterances were recorded on a SONY TCD-D3 Digital Audio Tape Walkman, using a SONY ECM-737 STEREO microphone or a SONY ECM-23FII Electret Condenser Microphone. Data was then transferred to 5-inch floppy discs through a Kay Elemetrics Visi-Pitch 6095/6097, and Nippon Electric Corporation (NEC) PC9801-RX, and the contents were printed using an NEC Printer.

The data was then visually inspected for differences in intensity, pitch, and length. One Japanese speaker and one American speaker were ultimately singled out for detailed study, and requested to read an additional list of sentences (see appendix). The final subjects were chosen because they displayed

employed in having the sentences read, also changed, over time, but most of the subjects in the preliminary survey were first asked to read a list of sentences without any directions, then handed the same list, but this time, with an emotion appended to each sentence. The subjects were asked to read the sentences, with the emotion listed beside each sentence. Some subjects commented that it was difficult to switch emotions frequently (for example, it was difficult to express anger, then surprise, then admiration, in succession) In the final phase of the research, two Americans and two Japanese were asked to read both sentence list 1 and sentence list 2 (Total 100 sentences). The subjects first read the sentences, with emotion, then were instructed to read the sentences, once in a monotonous tone, and then, immediately afterward, with emotion. Ultimately two final subjects were chosen for in-depth study.

A fourth drawback of the study is that several of the sentences are artificial because they are either: direct translations, or included to investigate certain aspects of American/Japanese intonational characteristics.

PRECAUTIONS

Several precautions were taken/learned to be taken when conducting this type of recording. A relaxed atmosphere is essential in drawing out the best response from the subjects, and small interruptions/too much emphasis on trying to obtain a noise-free recording, prove non-productive. The subjects respond far more enthusiastically, and produce the best recordings when encouraged, and tend to speak in louder tones, when the instructions given, in reading the sentences, are spoken in louder tones.

The subjects that give the best results are actor-type personalities, that have extensive experience in talking before other people, and the subjects who show the least emotional expression, appear to be scholarly-type individuals.

It is necessary to obtain a relatively noise-free recording, and a studio-type room or a quiet room for recording/appropriate equipment are essential.

SENTENCE TYPES

Sentences are divided into the 16 types listed below:

An	= Anger	Mv	= Moved
Az	= Amazement	Nst	= Nostalgia
Alm	= Alarm	Of	= Offended
Con	= Condemnation	Rz	= Resignation
Dsc	= Discovery	Sp	= Surprise
Dscrg	= Discouraged	Spaz	= Surprise and amazement
Indf	= Indifference	St	= Satisfaction
Indfcon	= Indifference and condemnation	Thk	= Thankfulness

RESULTS

(1) The number of pitch movements

Both the Japanese subject and the American subject increase pitch height, when asked to read the sentences with emotion, but the Japanese decreases the number of pitch movements, while the number of pitch movements for the American subject remains relatively the same. For instance, the Japanese subject places pitch movements on the " なんて" (nante) (What...)(Fig.1 a), " 赤い" (akai) (red)(Fig.1 b), and the "花" (hana) (flower)(Fig.1 c), in the monotonous sentence(なんて美しい赤い花だろう!)(nante ɯtsɯ̈kɯɕi: akai hana daro:)(What a beautiful red flower!): thus there are three major rise-fall combinations, but in the same sentence, read with emotion, there is one major rise-fall at the beginning of the sentence, on the "なんて" (nante), and a gradual decline, thereafter. The same tendency is manifested in other sentences of the same type. The sentence is slightly artificial, to test whether the increase in the number of words in a single sentence affects the pitch. The American subject, on the other hand, has relatively flat pitch curves when reading sentences monotonously, but distinct pitch movements, when reading with emotion. The

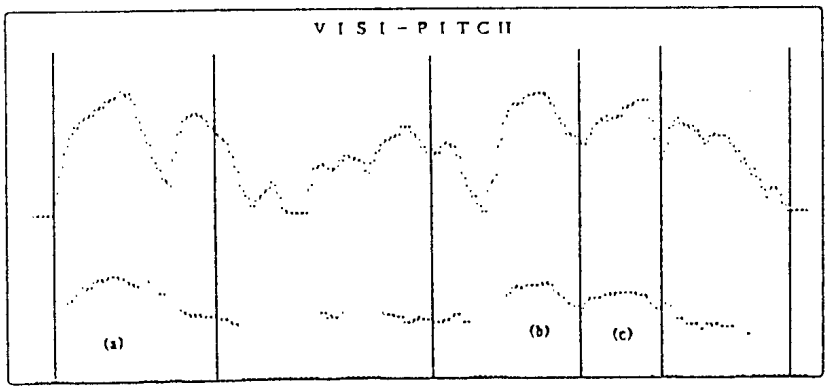

Fig.1

JIJINONO

1. なんて美しい赤い花だろう！ (間き、単心)

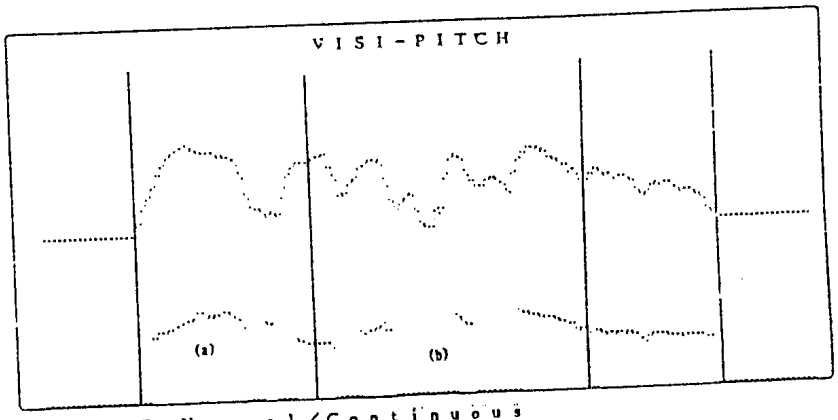

Fig. 2

AIJINONO

1. なんて美しいんだろう！ (間き、単心)

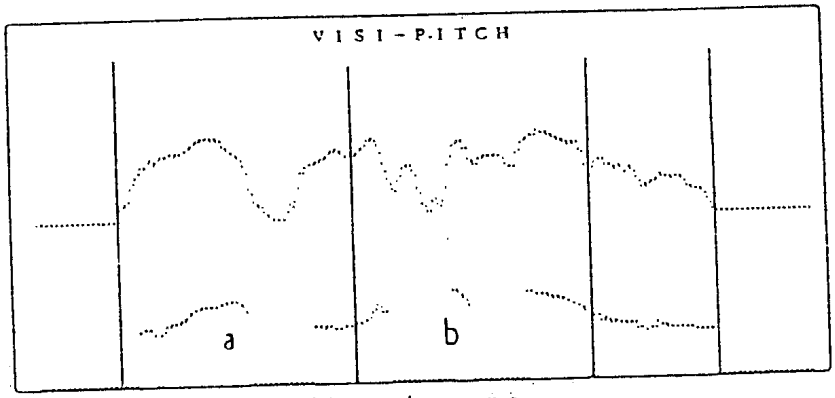

Fig.3

AIJINONO

1. なんて美しいんだろう！ (間き、単心)

number of pitch movements tends to remain about the same. For example, "なんて美しいんだろう!" (Nante ɯtsɯ̈kɯɕi:n daro:) (How beautiful) has about two small pitch movements on the "なんて" (Nante)(How) and the " 美しい" (ɯtsɯ̈kɯɕ i:) (beautiful), but the same sentence read with emotion has two rather large pitch movements, on the same words.

The decrease in the number of pitch curves by the Japanese speaker when reading with emotion, can be attributed to larger intonation groups. Since words in Japanese are spoken with distinct pitch movements, the pitch movements for the individual words form small intonation groups, when reading in a monotonous tone. However, when the Japanese subject reads with emotion, the small word-level pitch movements recede in importance, compared with the large group-level intonation movements. Thus, the Japanese subject reads emotional sentences with fewer, larger pitch movements compared with monotonous sentences. Meanwhile, since pitch movements do not seem to be as integral a part of a word, in English, as in Japanese, the American subject shows relatively few pitch movements when speaking in a monotonous tone (Fig.2 a & b), but uses larger pitch movements to emphasize important words ["なんて" (Nante) (How) and "美しい" (ɯtsɯ̈kɯɕi:) (beautiful)], when speaking with emotion (Fig.2 a & b).

(2) The location of the pitch movements

The words that have pitch movements are also different for the Japanese and American subjects. The Japanese informant generally has a pitch movement for almost every content word when reading monotonous sentences, but a large pitch rise at the beginning of emotional sentences, and a decline thereafter. Meanwhile, the American subject has relatively few pitch movements when reading monotonously, but uses pitch movement to emphasize important words in the sentence, when reading with emotion. The Japanese pitch, therefore seems to have both a lexical function (distinguishing words), and an emotional function (a rise and a fall over a large cluster of words - intonation group). The words which have a rise and fall in sentences read with emotion are not necessarily

important words, but rather, words that fall at the beginning, or at the end of intonation clusters. This means that the Japanese speaker's emotional pitch rise and pitch decline is determined by the place it occupies in the sentence. In other words, it is not the meaning of the words, or the contextual importance of the words that determine the overall emotional pitch movements, but rather, pitch movements occupy the crucial boundaries between large intonation groups. However, the American pitch, on the other hand, seems mainly to serve an emphatic function. Emotion and emphasis is expressed by placing pitch movements on important words - content words.

The Japanese speaker shows a steep pitch decline at the end of many utterances, and this signifies the end of an intonation group. The American informant, on the other hand, shows steep pitch declines at the end of sentences, only when that word is an essential element needed to convey important information. These traits are evident in English, where, for instance, in the sentence: "How beautiful you were, and how beautiful you are!", the Japanese subject shows a steep decline on the "are!" for both the monotonous (Fig.4 a) and emotional (Fig.5 b) sentences, but the American shows only a minute decline in the monotonous sentence (Fig.6 c), but a significant decline in the emotional sentence (Fig.7 d). This is important to clarify that the American speaker not only thinks that "you" was beautiful in the past, but also beautiful at present. Thus, the downward pitch movement is an integral part of the Japanese speaker's "are" (pitch is assigned to the word "are"), and this downward movement is accentuated when the overall intonation group is emphasized. However, the "are" is not assigned a pitch movement in the monotonous sentence by the American speaker, but rather the pitch is included for semantic purposes, in the emotional statement.

Thus, the Japanese subject can be said to have:

i. a steep pitch increase at the beginning of intonation groups, and a gradual or steep decline, sometime thereafter. The pitch change does not necessarily occur on an accented syllable, or an important word (for Japanese), while the

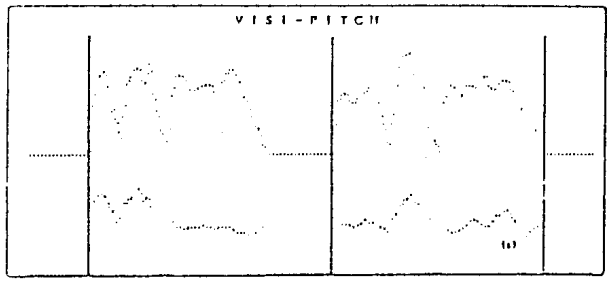

Fig.4

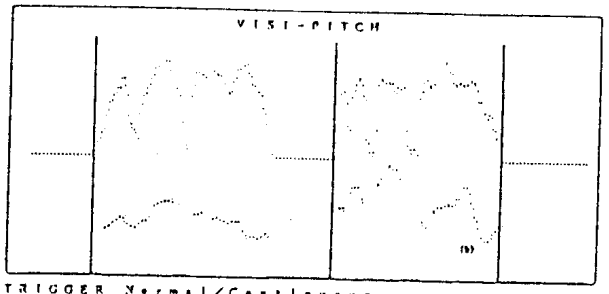

Fig.5

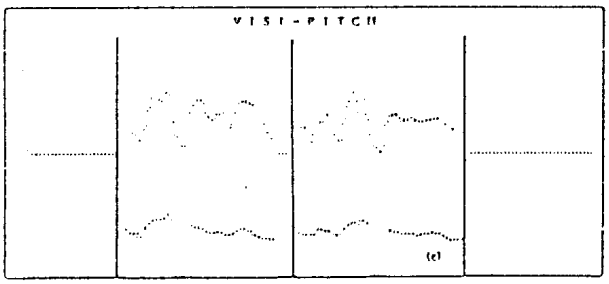

Fig.6

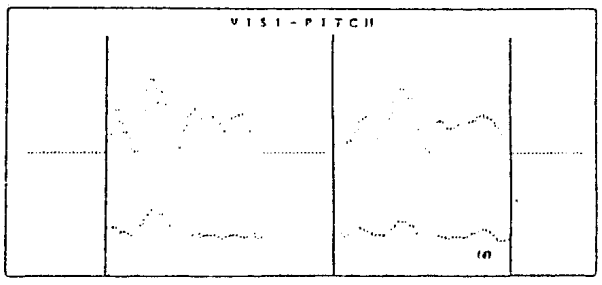

Fig.7

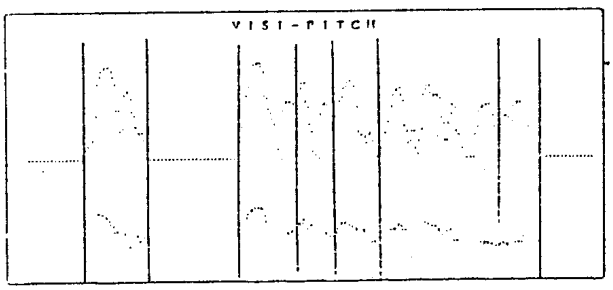

Fig.8

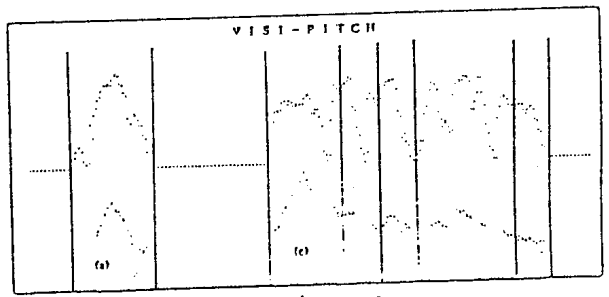

Fig.9

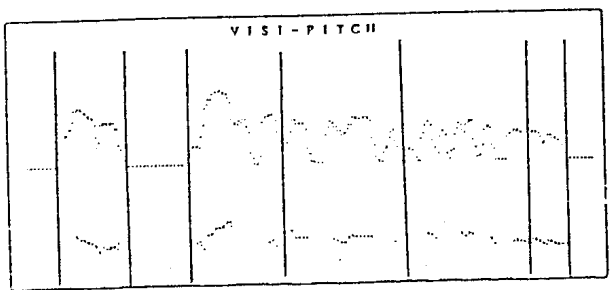

Fig.10

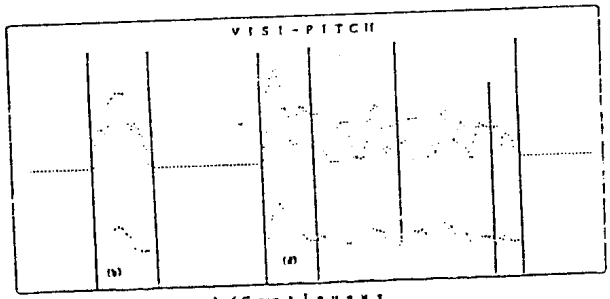

Fig.11

American subject can be said to have:

i. either a steep pitch increase or a steep pitch jump at the beginning of a prominent word, followed by an even more steep fall at the end of the word. The pitch movements lie on accented syllables, and are not related to the position of the word in the sentence (whether the word lies at the beginning, middle or end of an intonation group)

(3) The Direction of the pitch

The Japanese subject seems to place more emphasis on the pitch rise than the American subject, and the American seems to place more emphasis on the pitch fall, than the Japanese subject, to express emotion. For instance, in the sentence: "そうだ! どうしてもっと早く気が付かなかったんだろう!"(So:da! do:ɕite motᵑt'o hajakw̥kiga tsɯ̈kanakatᵑt'an daro:!), the Japanese subject's pitch both rises, then falls for the utterance, " そうだ!" (so:da) (That's it!)(Fig.9 a), while the American subject's pitch shows mainly a decline(Fig.11 b). Also, the Japanese subject has a quick rise and slow decline(Fig.9 c), while the American subject has a quick rise and quick decline(Fig.11 d) for " どうしてもっと早く気が付かなかったんだろう!" (do:ɕite motᵑt'o hajakw̥kiga tsɯ̈kanakatᵑt'an daro:!) (Why didn't I think of it before?!) Neither of these quick pitch rises are observed when the Japanese subject reads the sentence in a monotonous tone(Fig.8), but become evident when reading with emotion. Previous research has also revealed that it is the fall, not the rise in English, that determines which word is prominent, in a sentence. (Cruttenden 1986, p.51) Thus, the American shows only a moderate decline in " そうだ！" (so:da), and even rises, in " どうして" (do:ɕite), when speaking in a monotonous tone(Fig.10), but both words have a steep pitch decline when spoken with emotion(Fig.11).

(4) Pause

(a) Pause and Phrasing

The pause (pause as a proportion of total time taken to read the sentence) between utterances is significantly different for the American and Japanese speaker. The American speaker tends to shorten pauses more than the

Japanese, for emphasis, although this tendency depends on the type of sentence. For example, in the sentence: "Watch out! The dog's dangerous", the American speaker inserts a pause after "Watch out!", in the emotional sentence (Fig.15), thus increasing the overall pause from 2.05% (Fig.14 a) in the monotonous sentence, to 13.84% (Fig.15 b), in the emotional sentence. On the contrary, the Japanese subject decreases the overall pause from 25.02%(Fig.12 a) of the utterance in the monotonous statement, to 19.5%(Fig.13 b) in the emotional statement.

It is possible to interpret the American informant's pause as a re-phrasing of the sentence, to express emotion. The American speaker's monotonous sentence can be divided into three relatively equal portions - "Watch out", measuring 588 milliseconds, "the dog's", 551 ms, and "dangerous", 661 ms. The timing for the same items in the emotional sentence are: 644, 515, and 662 ms, respectively, but a 239-ms pause is inserted between the "Watch out!", and "the dog's". Thereby, the emotional sentence can be re-arranged into two large sections consisting of "Watch out!" + pause = 883 ms, and "The dog's dangerous! " = 1,177 ms.

Table 1

	monotonous	emotional	emotional re-grouped
Watch out	588 msec.	644 msec.	883 msec.
pause	000 msec.	239 msec.	
the dog's	551 msec.	515 msec.	1,177 msec.
dangerous	661 msec.	662 msec.	

Although this is not a perfect balance, the pause between the "Watch out!" and "The dog's dangerous!" serves to separate the two phrases, thus emphasizing both statements. First, the "Watch out!" draws the listener's attention to the speaker, and the explanation, "The dog's dangerous!" comes later.

This insertion of a pause/re-grouping of the phrasing is also evident in

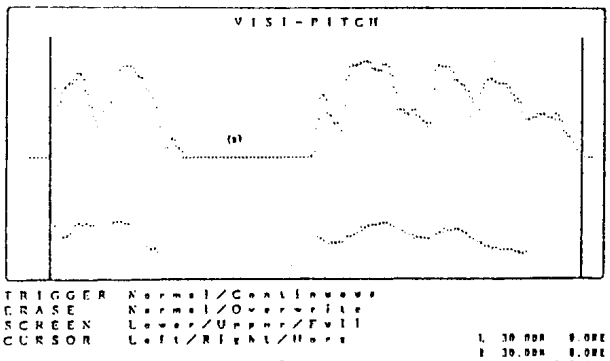

Fig.12

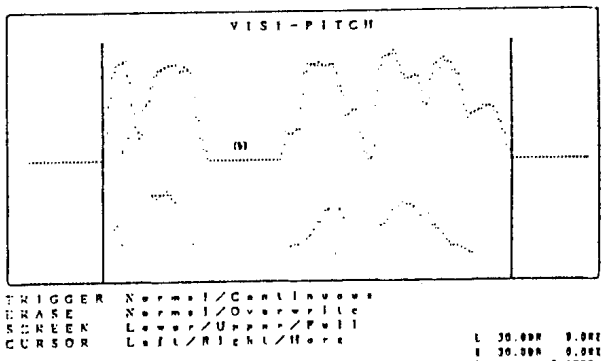

Fig.13

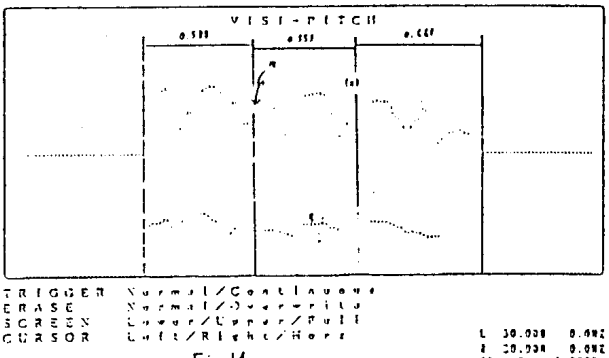

Fig.14

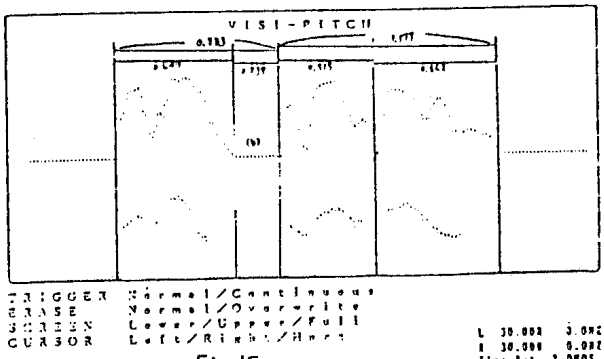

Fig.15

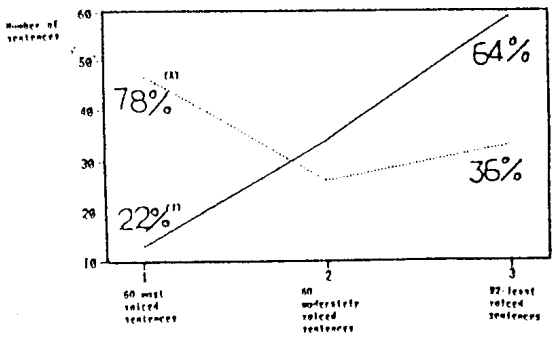

Fig.16

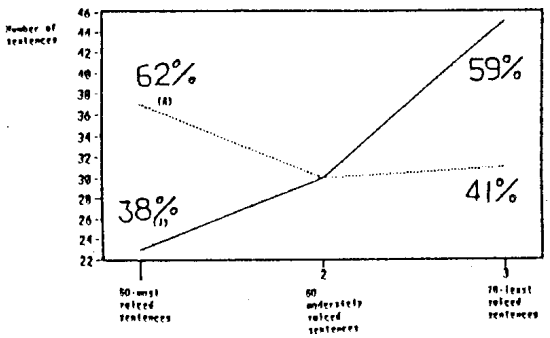

Fig.17

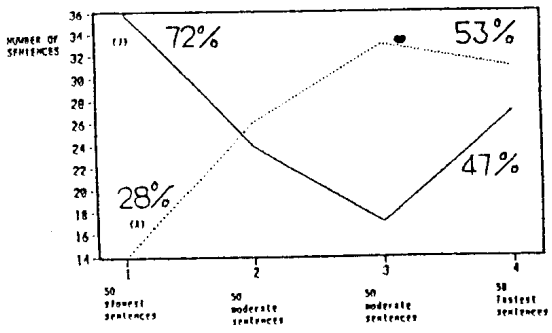

Fig.18

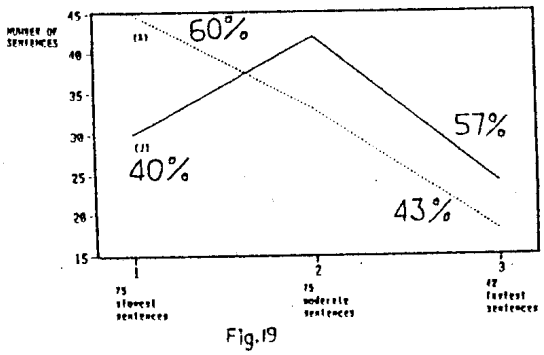

Fig.19

other sentences such as: "Watch out! He's dangerous!", and "Wow! What a beauty!"

Thus, the American subject inserts a pause in each of these sentence combinations to show emphasis, while in contrast, the Japanese informant shortens the pause between each of the same combinations. This may be attributed to the Japanese tendency to assign a steeper pitch rise/fall at the phrase boundaries to achieve emphasis, rather than the insertion/prolonging of a pause. The Japanese subject not only shortens pauses in "Watch out! The dog's dangerous!", and "Watch out! He's dangerous!", but shortens the pause more than any other element in the utterance. The sentence, "Wow! What a beauty!", meanwhile, is lengthened time-wise, on a whole, but the pause is completely eliminated between the "Wow!" and the "What a beauty!".

Therefore, the pause heightens the emotion, for the American informant, and serves to make each phrase stand out, whereas the pause serves to de-emphasize a phrase, for the Japanese informant.

(b) Regression Analysis and pause

Visipitch data: time, average frequency, standard deviation of the pitch, average intensity, standard deviation of the intensity, maximum pitch, minimum pitch, range of the pitch, percentage of voiced sounds, percentage of unvoiced sounds, percentage of pause, sentence type was input on two large Lotus 1-2-3 charts measuring 13 X 459 = 5967 cells (for the English), and 13 X 441 = 5733 cells (for the Japanese). Regression analysis was then applied to determine which factors are related for both the Japanese and American subjects when expressing emotion. Results are shown in Tables 1-4. The factors that are related for the Japanese subject, but not the American subject, and vice-versa are enclosed in the boxed areas.

First, the pause is related to numerous other factors for the Japanese subject, regardless of the sentence type. The Japanese subject decreases the average frequency, average intensity, maximum frequency, standard deviation of the frequency, and range of the frequency when increasing the pause to express

THE R-SQUARED VALUES FOR EACH COMBINATION OF VARIABLES IS SORTED IN DESCENDING ORDER (FROM LARGEST TO SMALLEST). THE LARGER THE R-SQUARED VALUE, THE STRONGER THE CORRELATION BETWEEN THE TWO VARIABLES. THE VARIABLES ARE CORRELATED IF THEY FALL ABOVE THE THICK HORIZONTAL LINE IN THE MIDDLE.

Table 1

AMERICAN SPEAKER CORRELATION SORT (ENG)

	R SQUARED	XCOEFFICNT
F0.RANGE MAX.F0	0.820	0.913
STD.DEV.F0 AVG.F0	0.667	0.473
F0.RANGE STD.DEV.F0	0.622	2.854
AVG.DB AVG.F0	0.499	0.095
MAX.F0 STD.DEV.F0	0.494	2.599
XUNVOICED XVOICED	0.471	-0.614
XPAUSE TIME	0.433	20.018
AVG.DB STD.DEV.F0	0.419	0.150
MAX.F0 AVG.F0	0.336	1.240
STD.DEV.DB AVG.DB	0.335	0.329
F0.RANGE AVG.DB	0.334	9.023
MAX.F0 AVG.DB	0.307	8.835
F0.RANGE AVG.F0	0.303	1.153
XPAUSE XVOICED	0.219	-0.360
STD.DEV.DB STD.DEV.F0	0.194	0.058
STD.DEV.DB AVG.F0	0.167	0.031
XUNVOICED MAX.F0	0.138	0.085
XVOICED TIME	0.118	-13.607
XPAUSE XUNVOICED	0.104	-0.261
XUNVOICED F0.RANGE	0.097	0.073
F0.RANGE STD.DEV.DB	0.095	0.190
XVOICED MAX.F0	0.085	-0.072
MAX.F0 STD.DEV.DB	0.077	0.186
AVG.DB TIME	0.065	2.563
XUNVOICED STD.DEV.F0	0.062	0.207
MIN.F0 MAX.F0	0.057	0.080
XVOICED F0.RANGE	0.042	-0.057
XPAUSE STD.DEV.DB	0.040	-1.061
MIN.F0 STD.DEV.F0	0.036	-0.256
XVOICED MIN.F0	0.036	-0.178
XUNVOICED MIN.F0	0.033	0.115
XUNVOICED TIME	0.030	-6.386
XVOICED STD.DEV.F0	0.028	-0.151
XUNVOICED STD.DEV.DB	0.025	1.022
XPAUSE AVG.F0	0.019	-0.056
F0.RANGE MIN.F0	0.016	-0.341
MIN.F0 AVG.F0	0.012	0.086
MAX.F0 TIME	0.012	17.613
F0.RANGE TIME	0.010	15.663
XUNVOICED AVG.F0	0.010	0.049
XPAUSE STD.DEV.F0	0.007	-0.057
XPAUSE F0.RANGE	0.007	-0.017
XUNVOICED AVG.DB	0.007	-0.312
AVG.F0 TIME	0.007	6.301
STD.DEV.F0 TIME	0.006	3.371
STD.DEV.DB TIME	0.005	-0.425
XPAUSE MAX.F0	0.005	-0.011
MIN.F0 STD.DEV.DB	0.005	-0.118
XPAUSE AVG.DB	0.003	-0.153
XVOICED AVG.DB	0.002	0.155
MIN.F0 AVG.DB	0.001	-0.190
XPAUSE MIN.F0	0.001	-0.015
MIN.F0 TIME	0.001	-1.282
XVOICED STD.DEV.DB	0.000	0.018
XVOICED AVG.F0	0.000	0.008

Table 2

JAPANESE SPEAKER CORRELATION SORT (ENG)

	R SQUARED	XCOEFFICNT
F0.RANGE MAX.F0	0.861	0.952
MAX.F0 AVG.F0	0.707	1.290
F0.RANGE STD.DEV.F0	0.658	2.780
MAX.F0 STD.DEV.F0	0.615	2.597
XUNVOICED XVOICED	0.572	-0.714
STD.DEV.F0 AVG.F0	0.569	0.349
F0.RANGE AVG.F0	0.517	1.111
AVG.DB AVG.F0	0.425	0.041
XPAUSE AVG.F0	0.316	-0.111
XPAUSE AVG.DB	0.313	-1.740
XUNVOICED STD.DEV.F0	0.310	0.333
STD.DEV.DB AVG.DB	0.282	0.228
XUNVOICED MAX.F0	0.273	0.096
XUNVOICED F0.RANGE	0.249	0.088
MAX.F0 AVG.DB	0.227	11.531
XUNVOICED AVG.F0	0.215	-0.130
XPAUSE MAX.F0	0.191	-0.056
XPAUSE STD.DEV.F0	0.175	-0.176
AVG.DB TIME	0.163	-1.819
XPAUSE TIME	0.157	8.621
XPAUSE XVOICED	0.150	-0.256
AVG.DB STD.DEV.F0	0.139	0.051
XPAUSE F0.RANGE	0.137	-0.018
STD.DEV.DB AVG.F0	0.137	-0.010
MAX.F0 STD.DEV.DB	0.137	1.117
F0.RANGE AVG.DB	0.133	1.180
MIN.F0 AVG.F0	0.106	-0.119
XUNVOICED STD.DEV.DB	0.094	-1.199
MIN.F0 AVG.DB	0.094	1.662
XPAUSE XUNVOICED	0.087	-0.206
AVG.F0 TIME	0.077	-30.695
STD.DEV.F0 TIME	0.074	-13.922
XUNVOICED TIME	0.071	-8.276
XVOICED AVG.DB	0.067	1.215
XVOICED STD.DEV.F0	0.065	-0.184
MAX.F0 TIME	0.058	-40.563
F0.RANGE STD.DEV.DB	0.058	1.056
XVOICED F0.RANGE	0.057	0.045
F0.RANGE MIN.F0	0.052	-0.656
F0.RANGE TIME	0.051	-39.580
XVOICED STD.DEV.DB	0.046	2.339
STD.DEV.DB TIME	0.046	-0.844
MIN.F0 STD.DEV.DB	0.043	1.172
XVOICED MAX.F0	0.042	-0.040
STD.DEV.DB STD.DEV.F0	0.028	-0.010
XPAUSE MIN.F0	0.024	-0.058
XVOICED MIN.F0	0.024	0.084
XPAUSE STD.DEV.DB	0.021	-1.060
XUNVOICED AVG.DB	0.019	0.604
MIN.F0 MAX.F0	0.014	0.042
MIN.F0 STD.DEV.F0	0.007	-0.099
XVOICED AVG.F0	0.003	-0.017
MIN.F0 TIME	0.001	2.174
XUNVOICED MIN.F0	0.001	-0.013
XVOICED TIME	0.000	-0.371

Table 3

JAPANESE SPEAKER CORRELATION JAPANESE

	R-SQUARED	XCOEFFICNT
F0.RANGE MAX.F0	0.829	0.957
F0.RANGE STD.DEV.F0	0.628	3.370
MAX.F0 STD.DEV.F0	0.605	3.148
XUNVOICED XVOICED	0.520	-0.550
XPAUSE XVOICED	0.420	-0.450
STD.DEV.F0 AVG.F0	0.359	0.257
MAX.F0 AVG.F0	0.333	1.000
MIN.F0 AVG.F0	0.232	0.391
XPAUSE MAX.F0	0.221	-0.091
AVG.DB STD.DEV.F0	0.215	0.183
MAX.F0 AVG.DB	0.212	4.705
XUNVOICED STD.DEV.DB	0.209	1.092
XUNVOICED AVG.F0	0.200	0.170
XUNVOICED TIME	0.196	-12.308
AVG.DB AVG.F0	0.189	0.071
XPAUSE F0.RANGE	0.156	-0.015
XVOICED STD.DEV.DB	0.154	-1.583
XPAUSE AVG.DB	0.145	-0.179
STD.DEV.DB TIME	0.138	-1.129
F0.RANGE AVG.DB	0.135	1.051
F0.RANGE AVG.F0	0.115	0.618
F0.RANGE MIN.F0	0.110	-0.751
XPAUSE STD.DEV.F0	0.101	-0.200
STD.DEV.DB AVG.DB	0.093	0.076
STD.DEV.DB AVG.F0	0.087	0.011
XUNVOICED MAX.F0	0.081	0.061
AVG.F0 TIME	0.083	-21.019
XPAUSE AVG.F0	0.066	-0.089
XPAUSE TIME	0.051	5.892
XUNVOICED MIN.F0	0.053	0.108
STD.DEV.DB STD.DEV.F0	0.045	0.021
XUNVOICED STD.DEV.F0	0.035	0.166
MAX.F0 STD.DEV.DB	0.035	1.555
XVOICED TIME	0.031	6.381
MIN.F0 AVG.DB	0.030	0.821
XUNVOICED F0.RANGE	0.028	0.015
XVOICED AVG.F0	0.027	-0.082
XVOICED AVG.DB	0.023	0.419
XVOICED F0.RANGE	0.022	0.010
XUNVOICED AVG.DB	0.021	0.318
F0.RANGE STD.DEV.DB	0.019	0.997
MAX.F0 TIME	0.018	-18.952
MIN.F0 STD.DEV.F0	0.017	-0.249
XPAUSE MIN.F0	0.014	-0.051
STD.DEV.F0 TIME	0.012	-3.423
MIN.F0 STD.DEV.DB	0.011	1.198
AVG.DB TIME	0.011	-1.311
F0.RANGE TIME	0.011	-14.265
XVOICED MAX.F0	0.011	0.030
XVOICED MIN.F0	0.009	-0.057
XVOICED STD.DEV.F0	0.006	0.093
MIN.F0 TIME	0.006	-4.481
MIN.F0 MAX.F0	0.006	0.035
XPAUSE XUNVOICED	0.004	-0.055
XPAUSE STD.DEV.DB	0.003	0.478

Table 4

AMERICAN SPEAKER CORRELATION JAPANESE

	R-SQUARED	XCOEFFICNT
XUNVOICED XVOICED	0.543	-0.542
MAX.F0 STD.DEV.F0	0.529	3.059
F0.RANGE MAX.F0	0.496	0.745
MIN.F0 AVG.F0	0.476	1.146
XPAUSE XVOICED	0.462	-0.457
MAX.F0 AVG.F0	0.432	1.376
STD.DEV.DB AVG.F0	0.392	0.051
F0.RANGE STD.DEV.F0	0.385	2.760
MAX.F0 STD.DEV.DB	0.377	15.671
STD.DEV.DB STD.DEV.F0	0.338	0.098
XPAUSE TIME	0.325	16.411
STD.DEV.F0 AVG.F0	0.270	0.258
F0.RANGE MIN.F0	0.199	-0.595
XUNVOICED MIN.F0	0.134	0.194
F0.RANGE STD.DEV.DB	0.116	1.180
MIN.F0 STD.DEV.DB	0.103	6.191
MIN.F0 MAX.F0	0.103	0.255
XVOICED F0.RANGE	0.102	0.113
XUNVOICED AVG.F0	0.100	0.279
XVOICED MIN.F0	0.095	-0.227
MIN.F0 AVG.DB	0.093	1.125
XUNVOICED F0.RANGE	0.088	0.118
XVOICED STD.DEV.F0	0.062	0.602
XVOICED TIME	0.058	-10.280
MAX.F0 AVG.DB	0.041	0.913
XUNVOICED TIME	0.040	-6.302
XPAUSE STD.DEV.F0	0.038	-0.309
XPAUSE AVG.DB	0.034	-0.320
AVG.DB STD.DEV.F0	0.030	0.167
XUNVOICED STD.DEV.F0	0.028	-0.191
AVG.DB AVG.F0	0.028	0.017
XPAUSE F0.RANGE	0.023	-0.055
XVOICED AVG.F0	0.022	-0.178
STD.DEV.F0 TIME	0.019	-2.452
F0.RANGE TIME	0.017	-10.329
AVG.DB TIME	0.017	-2.164
XPAUSE AVG.F0	0.015	-0.100
XVOICED AVG.DB	0.014	0.309
MIN.F0 TIME	0.012	6.571
XPAUSE MAX.F0	0.011	-0.030
F0.RANGE AVG.F0	0.011	0.230
XUNVOICED STD.DEV.DB	0.009	1.036
MIN.F0 STD.DEV.F0	0.008	0.298
XVOICED MAX.F0	0.008	0.051
XPAUSE MIN.F0	0.005	0.034
XVOICED STD.DEV.DB	0.003	-0.772
MAX.F0 TIME	0.003	-3.758
F0.RANGE AVG.DB	0.002	-0.212
XPAUSE STD.DEV.DB	0.000	-0.259
STD.DEV.DB AVG.DB	0.000	0.001
XUNVOICED MAX.F0	0.001	-0.010
AVG.F0 TIME	0.000	0.588
STD.DEV.DB TIME	0.000	-0.019
XUNVOICED AVG.DB	0.000	0.012
XPAUSE XUNVOICED	0.000	0.004

emotion. This indicates that the Japanese subject changes the pitch drastically when shortening the pause, to show emotion.

On the other hand, the American informant shows no correlation between the pause and the frequency/intensity overall, but does correlate these factors when expressing alarm in English, as the chart below indicates:

Table 5

Correlation between pause and other variables when the American speaker's expresses alarm

	R-squared	Sample Size	X-coefficient	R-value
Avg.Frequency	0.268	14	-0.284	0.518
Std.Dev.Frequ.	0.491	14	-0.573	0.700
Avg.Intensity	0.195	14	-1.746	0.442
Maximum Frequ.	0.316	14	-0.244	0.567
Range of Freq.	0.573	14	-0.226	0.757

The variable listed in the chart above are correlated with alarm when the R-value is above 0.661, thus the American subject increases pause when decreasing standard deviation of the frequency/range of the frequency for sentences read with alarm. Thus, the Japanese subject correlates the frequency and pause for all types of sentences, while the American only correlates frequency with the pause if the sentence expresses alarm.

(5) Voiced and unvoiced sounds

American speakers seem to unnecessarily prolong their vowels, when speaking in Japanese, and Japanese speakers seem to talk with a staccato rhythm, to some American speakers. These traits may be accounted for, by taking into consideration, the amount of voiced sounds spoken by both the American and Japanese subjects. The American subject has more voiced sounds for both the Japanese sentences and the English sentences, than the Japanese subject. Figure

16 shows that the American subject accounts for 78% of the 60 most-voiced sentences/utterances in English, but only 36% of the 92 least-voiced sentences/utterances. Likewise, Figure 17 indicates that the American subject accounts for 62% of the 60 most-voiced sentences/utterances in Japanese, but only 41% of the 76 least-voiced sentences/utterances.

This may also account for the significant increase in voiced sounds the Japanese subject manifests when applying emotion to a statement, and the relatively small increase in the American subject, for the same factor. Fig.21 indicates that the Japanese increases voiced sounds as a percentage of total sounds by a significant margin (by over 10%) for 29% and 23% of the Japanese and English sentences respectively, when expressing emotion, while the American subject does so for only 4% and 9% of the sentences, respectively. It is probable that the Japanese subject is able to increase the amount of voiced sounds to such a great extent, because he has fewer voiced sounds than the American, to begin with, thus allowing greater room for increase.

(6) Time

The American subject takes longer to read the Japanese sentences than the Japanese subject, and the Japanese subject takes longer to read the English sentences than the American subject, which may also partially account for some degree of unnaturalness when perceived by the native speaker. The Japanese subject accounts for 72% of the slowest sentences, but only 47% of the 58 fastest sentences when speaking in English (Fig.18). On the other hand, the American subject accounts for 60% of the slowest sentences/utterances, but only 43% of the fastest sentences/utterances in Japanese (Fig.19)

(7) Lengthening to acheive emphasis

Data was sorted in order of all variables, and it was discovered that the Japanese subject lengthens 68% and shortens 32% of all sentences, to express emotion (Fig.20).

(8) Standard deviation of the frequency, Average frequency, and Minimum Frequency

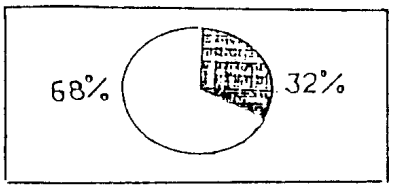

Fig.20

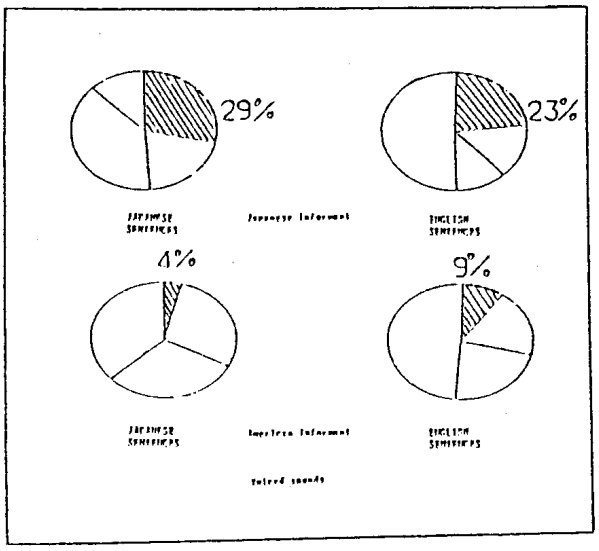

Fig. 21

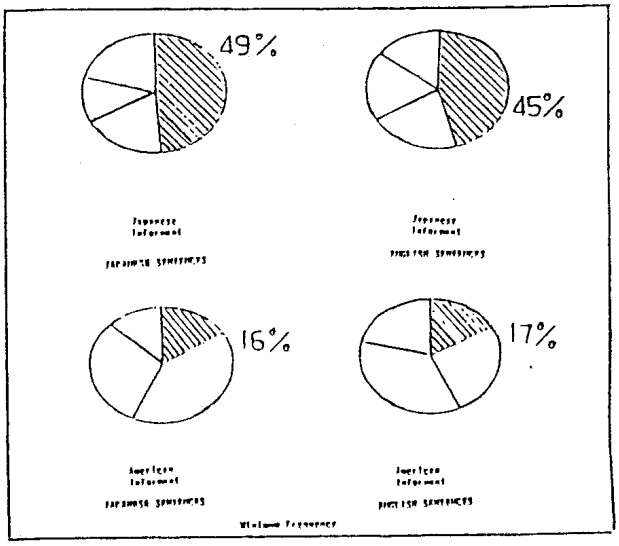

Fig.23

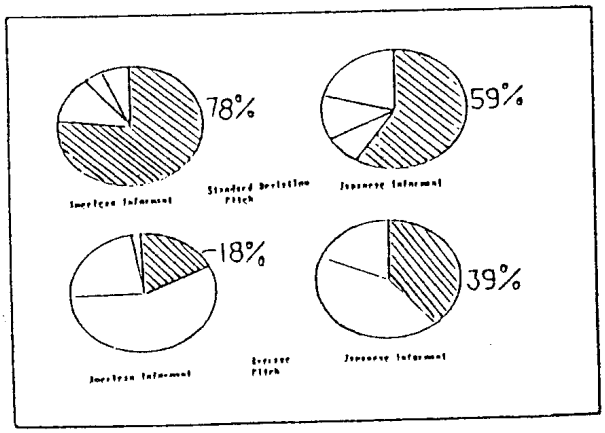

Fig.22

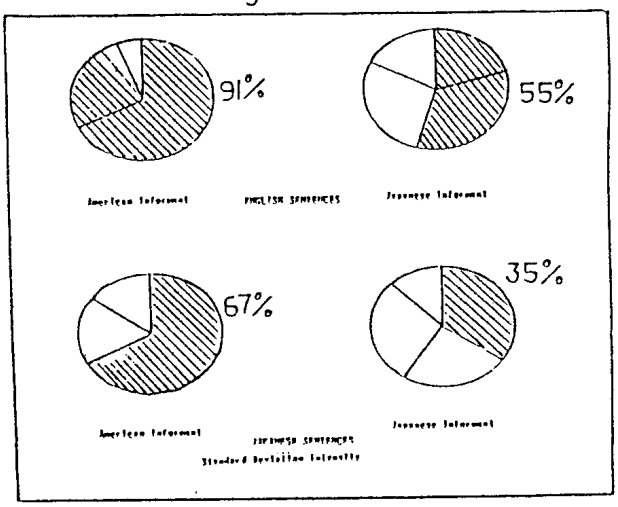

Fig. 24

The Japanese subject raises the frequency of the whole utterance (the minimum frequency as well as the maximum frequency), while the American tends to raise only the maximum frequency, to express emotion. The American subject tends to increase the standard deviation of the frequency (raises it 78% of the time) more than the Japanese (raises it 59% of the time), whereas the Japanese tends to increase the average frequency (raises it 39% of the time) more than the American (raises it 18% of the time), to express emotion (Fig.22). Furthermore, Fig.23 shows that the Japanese subject increases the minimum frequency (the lowest pitch) for 49% and 45% of the Japanese and English sentences, respectively, to express emotion, whereas the American does so for only 16% and 17%, respectively.

(9) Standard Deviation of the Intensity

Although the intensity is one of the less prominent factors able to be recognized by the listener, increase of the standard deviation of the intensity is significantly greater for the American speaker (increases intensity in 91% and 67% of the English and Japanese sentences, respectively, when applying emotion) than the Japanese (increases 55% and 35%, respectively)(Fig.24).

CONCLUSION

Thus, it is possible to see that a great range of factors lead to the sense of unnaturalness a speaker demonstrates when talking in a foreign language, and although such factors are difficult to filter out through simple observation, experimental analysis is able to accurately describe many of the differences that exist between a native language speaker, and a foreign speaker. By factoring out these differences, it is possible for a foreign language speaker to become aware of potential areas of improvement, and through training, acquire more of the features that the native speaker would employ in speech.

REFERENCES

Cruttenden, Alan (1986) Intonation , Cambridge University Press

Richards, Jack; Platt, John; Weber, Heidi (1985) Longman Dictionary of Applied Linguistics

Sentence List 2 Japanese

1 ベッドから火が出ている！ (危機感)

2 助けて！ 焼け死んじゃう！ (危機感)

3 火事だ！ 妻がまだ中にいるんです！ 助けてください！ (危機感)

4 あっ！ 危ない！ かみつくよ！ (犬を見て) (危機感)

5 あいつは危ないんだ！ 近づくな！ (危機感)

6 すごい！かっこいいな！ (新しい車を見て感動)

7 浩司: あっ！しまった！ 忘れちゃった！ (大変なことに気付き、パニック)

真 : 何を？！ (何を忘れたんだ？)

浩司: 切符！ (パニック)

真 : えっ！じゃ、間に合わないよ！ あと一時間で飛行機が出発するんだよ！(切迫)

8 和夫: おい！そんなことやっちゃだめだよ！ (非難)

満 : どうして？ (無関心)

和夫: 違法じゃないか！ (非難)

満 : 誰も見てやしないよ。 (どうでもいいじゃないか)

和夫: おまえはどうかしてるよ！ (非難、面くらい、失望)

満 : うるさいな！ ほっといてよ！ (いらだち)

9 これ、なんてひどいんだろう！ (失望)

10 これ、なんて素晴らしいんだろう！ (歓喜)

11 誰がこんなことを予想しただろう？！ (驚き)

12 もう、二度と話さない！ (怒り)

Sentence List 2 English

1. Help! The bed's on fire! (alarm)
2. Help! The bed is on fire! (alarm)
3. Help! They're on fire! (alarm)
4. Help! They are on fire! (alarm)
5. Help! The house is on fire, and my wife's inside! (alarm)
6. Help! It's on fire, and she's inside! (alarm)
7. Watch out! The dog's dangerous! (alarm)
8. Watch out! He's dangerous! (alarm)
9. Wow! What a beauty! (admiration at seeing a new car)
10. Bill: "Oh no! I forgot it!" (panic)
    Mark: "What?!" (What did you forget?)
    Bill: "My ticket!" (panic)
    Mark: "What?! We won't make it, then! The plane leaves in an hour!" (urgency)
11. Paul: "You can't do that!" (condemnation)
    Tom: "Why not?" (indifferent)
    Paul: "It's illegal!" (condemnation)
    Tom: "No one'll find out." (who cares?)
    Paul: "You're crazy!" (insistent, bewildered)
    Tom: "Get away!" (irritated... leave me alone)
12. What absolute hell this all is! (discouraged)
13. What absolute bliss, this is! (elated)
14. Who could have ever imagined such a thing! (surprised)
15. I'll never talk to you again! (angry)

1. bet°t'o kara çi ga dete irɯ
2. tasɯkete jakecindzaɯ
3. kadzida tsɯmaga mada nakaɲi irɯn°desɯ tasɯkete kɯdasai
4. a? abɯnai kamitsɯkɯjo
5. aitsɯwa abɯnain°da tcika°zɯkɯna
6. sɯgoi kak°k'oi:na
7. a? cimat°t'a ɰasɯretcat°t'a
   naɲi wo
   kip°p'ɯ
   e? dʑa maɲiawanaijo
   ato itcidʑikande çiko:ki ga cɯp°p'atsɯ sɯrɯn dajo
8. oi sonnakoto jat°t'ca damedajo
   do:cite
   iho:dʑanaika
   daremo mitejanai jo
   omaewa do:ka citerɯjo
   ɯrɯsaina hot°t'oite jo
9. kore nante çidoin daro:
10. kore nante sɯbaracin daro:
11. darega kon°nakotoo joso: cita daro:
12. mo: ɲidoto hanasanai