PTC ... Orig ... Recd 1/05/95

X

Tenth Quarterly Progress Report

November 1, 1994, through January 31, 1995

NIH Contract N01-DC-2-2401

Speech Processors for Auditory Prostheses

Prepared by

Dewey T. Lawson, Blake S. Wilson and Mariangeli Zerbi

Center for Auditory Prosthesis Research
Research Triangle Institute
Research Triangle Park, NC 27709

This QPR is being sent to you before it has been reviewed by the staff of the Neural Prosthesis Program

CONTENTS

I. Introduction 3

II. A Channel-Specific Tool for Analysis of Consonant Confusion Matrices. 4

III. Plans for the Next Quarter 12

Appendix 1: Token-Specific Lists of Channel-Specific Cues. 13

I. Introduction

The purpose of this project is to design and evaluate speech processors for implantable auditory prostheses. Ideally, the processors will extract (or preserve) from speech those parameters that are essential for intelligibility and then appropriately represent those parameters for electrical stimulation of the auditory nerve or central auditory structures. Work in the present quarter included the following:

1. Initial studies with the third of six patients in the Nucleus percutaneous series (NP3, January 9-20) and continued studies with the second patient in the series (NP2, January 30 to February 10). Studies with NP3 included evaluations of *continuous interleaved sampling* (CIS) and *spectral peak* (SPEAK) processing strategies. Studies with NP2 included repeated measures with the SPEAK strategy, detailed evaluation of CIS processors using more than six channels, and measures of intracochlear evoked potentials.
2. Incorporation of a high-rate, high-resolution A/D converter for the laboratory speech processor system (Burr Brown DEM-PCM1760).
3. Development and initial construction of new current sources for the laboratory system, capable of presenting short-duration (e.g., 10 μs/phase) pulses with high fidelity and minimal crosstalk among separate channels. The new current sources will support an evaluation of high carrier rates in CIS processors, as proposed in Quarterly Progress Report 9 for this project.
4. Evaluation of upward and downward extensions of the frequency range spanned by CIS channels, in studies beginning with subjects NP2 and NP3.
5. Continued analysis of prior speech reception and evoked potential data, and continued preparation of manuscripts for publication.

In this report we describe the development and initial application of a new tool to analyze results from consonant identification tests. This tool, called "IngImag," is designed to indicate differences in the ability of various processor channels to convey anticipated consonant cues. The resulting diagnostic information can provide guidance for adjustment of channel gains and/or mapping functions to improve speech reception scores for individual subjects.

Results from other studies and activities, as indicated in the list above, will be presented in future reports.

II. A Channel-Specific Tool for Analysis of Consonant Confusion Matrices

Among the many possible parametric manipulations available within the basic CIS processor approach, channel-by-channel adjustments of relative gains and mapping laws are likely to be among those with the greatest potential for improving speech reception performance. It is not difficult, for instance, to imagine patients with varying patterns of neural survival along their cochleae resulting in quite different optimal mapping laws for different electrodes. Our exploration of such possibilities has been very limited to date, however, because of the enormous commitment of time required for a systematic study. In the absence of any basis for assuming the effects of such manipulations to be orthogonal, one must expect a *geometric* increase in the time required to map the performance effects of a given type of adjustment if it is to be made on a channel-by-channel basis.

In the course of our many CIS processor studies, however, we have come across instances in which the failure of information transmission in a particular channel was unmistakable. In such cases we have achieved substantial improvements in overall performance from an adjustment to the gain of a single channel. Undertaking channel-by-channel adjustments *only on a limited number of channels known to have particular potential for improvement* would involve a much more reasonable committment of time -- both now in the laboratory and eventually in clinical fittings.

To this end we have designed a channel-specific tool for analysis of consonant confusion matrices.

Given the bandpass filters selecting signals for each channel in a particular processor and a specific set of consonant speech test tokens, we prepare a list of those channels expected to convey a salient cue in support of the identification of each token. The same set of tokens is used in an identification test to produce a consonant confusion matrix. Each off-diagonal element in such a matrix represents a particular incorrect identification, for which we can compare our list of salient channel cues for the token that was presented to our list for the token that was (incorrectly) identified. In doing this we can note (1) channels in which salient cues should have been present for the presented token but not for the token identified, and (2) channels in which salient cues would be expected in the case of the identified token but were not present in the presented token. As a sort of shorthand, we refer to these two cases as channel cues seemingly *ignored* by the subject, and channel cues seemingly *imagined* by the subject in arriving at his or her response. Accordingly, the computer program that performs this type of analysis has been dubbed **IgnImag**. Its analysis criteria can be summarized in the following decision table:

presented token cue expected in this channel?	response token cue expected in this channel? yes	no
yes	consistent	**ignored cue**
no	**imagined cue**	consistent

Clearly, the frequent occurrence of *ignored* cues in a particular channel could signal a general lack of salience for that channel. On the other hand, a pattern of frequent errors that involve *imagined* cues in a particular channel could also indicate such a lack of salience, reflected in an increased willingness to entertain responses consistent with the presence of unheard cues there. A pattern of frequent *imagined* cue errors also could be caused by a channel with too high a gain setting, producing extraneous cues.

Channel-specific cues for consonant identification include both relatively large and relatively small amplitude signals, transitions in preceding vowel signals, and in some cases notably small levels in particular channels may be significant cues as well. Rather than attach a numerical weight to each cue in our analysis, we have constructed more than one cue list allowing us to consider, for instance, ignored and imagined "major consonant cues" (large signals) first, and only when those cues are being conveyed reliably turn to "minor consonant cues" (small signals) and such other cues as characteristic silences and preceding vowel transitions.

As an illustration of how lists of channel-specific cues for these four categories are constructed, each of the following four figures displays data for a specific medial consonant token uttered by a male talker. The top trace in each case is the raw audio signal input to a specific CIS processor. Below are envelope signals for each of the six channels of that processor in order from the lowest frequency channel (number 1) to the highest (number 6). The first example (Fig. 1) is the token "asha", illustrating major consonant cues in channels 5 and 6, and characteristic silences in channels 1 and 2. In the second example (Fig. 2), "aha", minor consonant cues have been identified in channels 2, 3, and 4, and a silence in channel 1 also noted as a potential cue. "Awa" in Fig. 3 evidences silences in channels 3, 4, 5, and 6 and major consonant cues in the remaining two channels. In the final example, "aya" (Fig. 4) a characteristic transition from the preceding vowel has been listed as a potential cue in channel 1, along with major consonant signals in channels 1 and 5, minor consonant signals in channels 4 and 6, and a notable silence in channel 3.

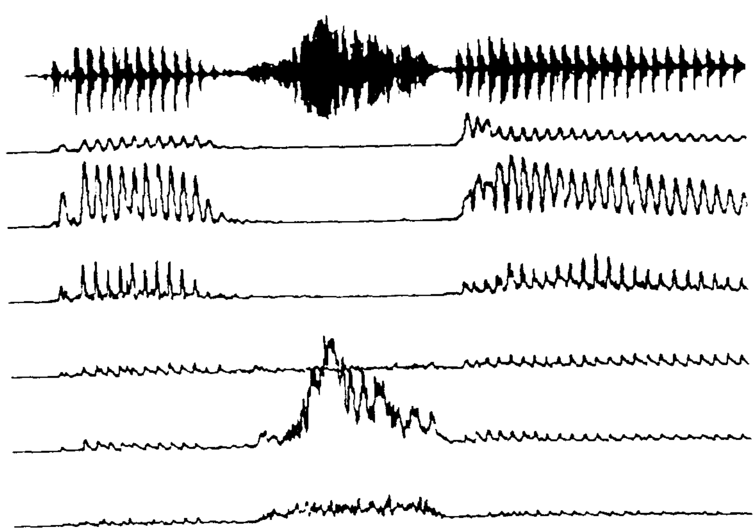

Fig. 1 Six Channel "asha", Male Talker

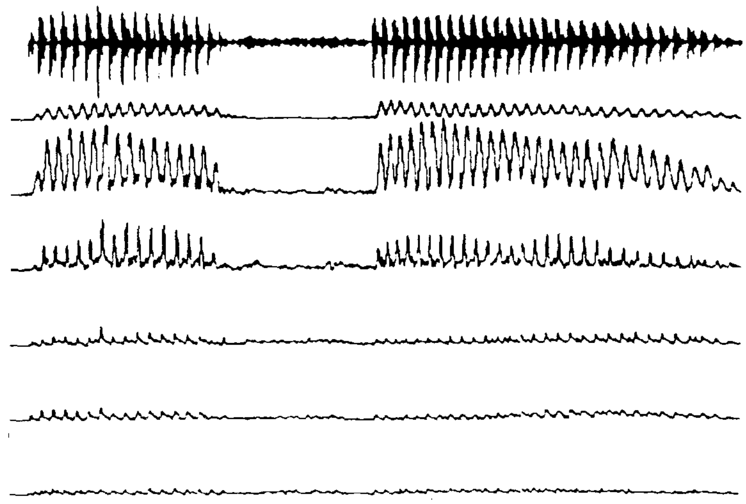

Fig. 2: Six Channel "aha", Male Talker

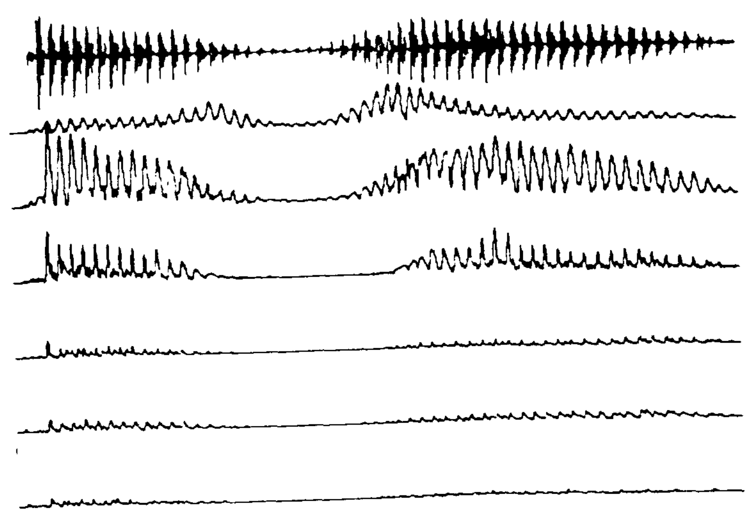

Fig. 3: Six Channel "awa", Male Talker

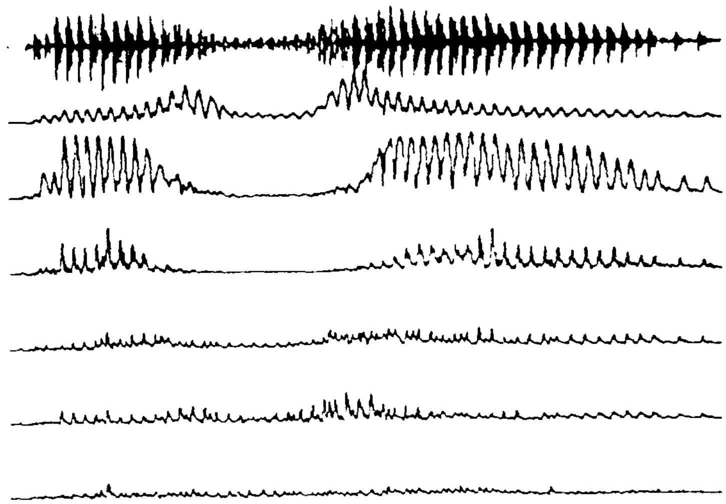

Fig. 4: Six Channel "aya", Male Talker

The analysis program **IgnImag** accepts keyboard entry of identification codes for a particular patient and a particular speech processor. All available archives are searched for any corresponding 16- or 24-consonant identification test results and a table of such tests is displayed, including relevant testing conditions, times, dates, and sequence numbers. Any subset of the available test results listed in this table may be selected for analysis.

IgnImag then displays the aggregate *ignored* and *imagined* cues for each channel, summed over the confusion matrices for the selected consonant identification test(s), both as absolute numbers and as percentages of the number of channel specific cues supplied in the presented tokens. Percentages are displayed not only for ignored and imagined cues in each category and each processor channel, but also summed across channels in each category and across categories in each channel.

The following results for a recent patient illustrate the potential utility of **IgnImag**. The underlying consonant confusion data were obtained for ten sets of 16 consonant tokens uttered by a male talker and presented as sound alone. The patient was using a 6 channel CIS processor, percutaneously stimulating six electrodes from a 22 electrode intracochlear array with percutaneous access.

Medial Consonant Confusion Tests: Patient NP-2, Processor 2

Item	Seq	Cons	Gend	Vis	Rand	Date	Time
2	32	16	m		4	940913	1346
3	33	16	m		5	940913	1355

Channel Cue Errors

Ignored

	Tot as %	% of Supplied by Chan						Count by Chan					
	of Suppl	1	2	3	4	5	6	1	2	3	4	5	6
Maj Cons	7	0	10	15	0	20	0	0	1	3	0	8	0
Min Cons	9	0	0	7	6	16	13	0	0	2	3	13	9
Vow Trans	6	8	0	0	0	0	0	3	0	0	0	0	0
Silences	2	4	0	1	1	1	1	4	0	1	1	1	1
Aggregate	5	4	1	4	2	12	5	7	1	6	4	22	10

Imagined

	Tot as %	% of Supplied by Chan						Count by Chan					
	of Suppl	1	2	3	4	5	6	1	2	3	4	5	6
Maj Cons	14	25	0	0	0	30	18	5	0	0	0	12	9
Min Cons	4	0	10	10	4	2	3	0	3	3	2	2	2
Vow Trans	14	15	10	0	0	0	0	6	1	0	0	0	0
Silences	4	3	2	7	4	4	4	3	2	6	3	3	3
Aggregate	6	7	4	6	3	9	7	14	6	9	5	17	14

While the overall performance of this patient with this processor was good (86 percent overall information transmission), **IgnImag** results indicated that consonant identification cues were least salient in channel 5, where 20 percent of major consonant cues and 16 percent of minor consonant cues were ignored (i.e., not reflected in the patient's responses). Twice as many major consonant cues were imagined as were ignored in this case, with such errors again concentrated in channel 5. This indication was consistent with the principal information transmission deficit of this processor, which was a 65 percent score in place of articulation, and focussed our attention on manipulations involving channel 5 for further improvements in 6 channel processor performance for this patient.

	--------------------	--------------------	Percent Information Transmission	--------------------	--------------------		
Processor	Correct	Overall	Voicing	Envelope	Frication	Place	Duration	Nasality
2	81	86	90	95	100	65	76	100

The crucial role of channel 5 in limiting the performance of this processor was dramatically confirmed when a similar five channel processor using the same electrodes except for the one previously assigned to channel 5 was found to support significantly better consonant recognition.

	--------------------	--------------------	Percent Information Transmission	--------------------	--------------------		
Processor	Correct	Overall	Voicing	Envelope	Frication	Place	Duration	Nasality
14	93	94	95	97	100	92	100	100

In a patient with only six available intracochlear electrodes, the alternatives to reducing the number of channels would have consisted of measures designed to improve the salience of cues transmitted via the problematic channel, e.g., adjustments in loudness balance with respect to the other channels. In this case, however, large improvements in the performance of a six channel processor were realized simply by substituting a different nearby electrode for the one previously assigned to channel 5 (electrodes 1, 5, 9, 13, 15, and 21 as numbered from the apical end, rather than 1, 5, 9, 13, 17, and 21).

	--------------------	--------------------	Percent Information Transmission	--------------------	--------------------		
Processor	Correct	Overall	Voicing	Envelope	Frication	Place	Duration	Nasality
41	93	93	95	97	100	83	94	100

The equivalent **IgnImag** analysis for this identical six channel processor with a different electrode assigned to channel 5 was as follows:

Medial Consonant Confusion Tests: Patient NP-2, Processor 41

Item	Seq	Cons	Gend	Vis	Rand	Date	Time
1	95	16	m		2	940922	0848
2	96	16	m		3	940922	0856

Channel Cue Errors

Ignored

	Tot as %	% of Supplied by Chan						Count by Chan					
	of Suppl	1	2	3	4	5	6	1	2	3	4	5	6
Maj Cons	1	0	0	5	0	0	0	0	0	1	0	0	0
Min Cons	4	0	0	0	2	9	7	0	0	0	1	7	5
Vow Trans	6	8	0	0	0	0	0	3	0	0	0	0	0
Silences	0	1	0	1	0	0	0	1	0	1	0	0	0
Aggregate	2	2	0	1	1	4	3	4	0	2	1	7	5

Imagined

	Tot as %	% of Supplied by Chan						Count by Chan					
	of Suppl	1	2	3	4	5	6	1	2	3	4	5	6
Maj Cons	5	0	0	0	0	10	10	0	0	0	0	4	5
Min Cons	2	0	3	3	0	5	0	0	1	1	0	4	0
Vow Trans	2	2	0	0	0	0	0	1	0	0	0	0	0
Silences	2	3	0	0	4	4	4	3	0	0	3	3	3
Aggregate	3	2	1	1	2	6	4	4	1	1	3	11	8

It should be noted that there was no indication of a significant difference between electrodes 15 and 17 in the current amplitudes required to achieve threshold and most comfortable loudness levels for the pulsatile stimuli of these processors:

Electrode	Current Amplitudes (ua) Threshold	MCL	Dynamic Range (dB)
1	187	390	6.4
5	167	417	7.9
9	139	402	9.2
13	103	323	9.9
15	102	322	9.9
17	103	314	9.7
21	161	337	6.4

While limited for the time being to analysis of six channel processors, **Ignimag** nevertheless proved useful in identifying (across multiple groups of six) a subset of electrodes that supported even better performance with an 11 channel CIS processor.

Processor	Correct	-------------------Percent Information Transmission-------------------- Overall	Voicing	Envelope	Frication	Place	Duration	Nasality
66	98	98	100	100	100	97	100	100

We plan systematically to test and calibrate this type of analysis in a series of studies with a research subject and CIS processor known to achieve high levels of accuracy on underlying 24 consonant identification tests. Intentional deficits of various types and degree will be produced in particular channels and the resulting effects on the **IgnImag** results observed. Among the fruits of such tests may be the identification of more efficient but equally sensitive subsets of consonant test tokens.

Preliminary token-specific lists of channel-specific cues have been completed in support of **IgnImag** analysis for a standard set of six channel frequency bands (extending from 350 to 5500 Hz). The lists were derived from our 24 medial consonant token set in /aCa/ context based on IOWA laserdisc recordings, for both male and female speakers. These lists are shown in Appendix 1 to this report. The next such lists to be constructed in support of our ongoing research will be for 11 channel processors and 6 channel processors covering a wider frequency range (350 to 9500 Hz).

III. Plans for the Next Quarter

Our plans for the next quarter include the following:

1. A meeting at RTI with Jim Patrick, Ron West and Jim Heller, to discuss progress and plans for studies involving patients with percutaneous access to the Nucleus electrode array (Wilson and Lawson, February 2).
2. Presentation of project results in one invited lecture and one poster at the annual *Midwinter Meeting of the Association for Research in Otolaryngology* (Wilson and Finley, February 5-9).
3. A meeting in Boston with Don Eddington, Marco Pelizzone, Jacques François and others, to discuss progress and plans for the continued joint development of a highly flexible and portable speech processor system (Zerbi, February 12-15).
4. Presentation of project results at Otolaryngology Grand Rounds, Duke University Medical Center (Lawson, February 15).
5. Continued studies with Nucleus percutaneous subjects NP2 (January 30 to February 10) and NP1 (March 13-24), including for each of the subjects repeated measures with the SPEAK strategy, detailed evaluation of CIS processors using more than six channels, and measures of intracochlear evoked potentials.
6. Continued studies with Ineraid subject SR2, including electrophysiological and psychophysical measures of forward masking for a wide range of maskers with various burst durations, pulse amplitudes, and pulse rates (February 20 to March 6).
7. A trip to the University of Innsbruck, at the invitation of the University, to participate in research there on the fitting and evaluation of CIS processors (Lawson, March 5-10).
8. Continued studies with Ineraid subject SR3, including recordings of intracochlear evoked potentials for a wide variety of stimuli, electrophysiological and psychophysical measures of forward masking, and initiation of chronic studies with a portable CIS processor, as implemented in the new processor developed at the University of Innsbruck with our help (April 3-13).
9. Presentation of project results at the *IIIrd International Congress on Cochlear Implant*, to be held in Paris, France (Wilson, April 27-29).
10. Continued analysis of speech reception and evoked potential data from prior studies, and continued preparation of manuscripts for publication.

Appendix 1

Token-Specific Lists of Channel-Specific Cues
24 Medial Consonant Token Set
Six Channel Frequency Bands from 350 to 5500 Hz

Male Talker

Token	Maj Cons						Min Cons						Vow Trans						Silence					
	--channel--						--channel--						--channel--						--channel--					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
m	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0
v	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	1	0	0	0
s	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
z	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0
sh	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
th(v)	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0
p	0	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	1	1
b	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1	1	1
t	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
d	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	1	1	1	1	1	1
g	0	0	0	1	0	0	1	1	1	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1
k	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1
j	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1
l	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
r	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1
w	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
y	1	0	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0
ng	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
h	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
zh	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0
th(u)	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0
ch	0	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	1

Female Talker

----------------Types of Cues------------------

Token	Maj Cons						Min Cons						Vow Trans						Silence					
	--channel--						--channel--						--channel--						--channel--					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
m	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
v	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	0
z	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	0	0	1	1	0	0	1	0
sh	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
th(v)	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1
p	0	0	0	0	1	0	1	1	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	1
b	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	1	1	1
t	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	1	1
d	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1
g	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	1
k	0	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	1	1
j	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1
l	1	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
r	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1
w	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1
y	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0
ng	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
h	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0
zh	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0
th(u)	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0
ch	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0